*[handwritten: Ulm, J. 09/899295 Seq ID 2 w/interf Page 1]*

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:  June 30, 2003, 16:16:45 ; Search time 311 Seconds
         (without alignments)
         744.242 Million cell updates/sec

Title:          US-09-899-295-2
Perfect score:  1860
Sequence:       1 MTLESIMACCLSEEAKEARR.........VFAAVKDTILQLNLKEYNLV 359

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       4569144 seqs, 64473110 residues

Total number of hits satisfying chosen parameters:   4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Pending_Patents_AA_Main:*
    1:/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
    2:/cgn2_6/ptodata/1/paa/US06_COMB.pep:*
    3:/cgn2_6/ptodata/1/paa/US07_COMB.pep:*
    4:/cgn2_6/ptodata/1/paa/US080_COMB.pep:*
    5:/cgn2_6/ptodata/1/paa/US081_COMB.pep:*
    6:/cgn2_6/ptodata/1/paa/US082_COMB.pep:*
    7:/cgn2_6/ptodata/1/paa/US083_COMB.pep:*
    8:/cgn2_6/ptodata/1/paa/US084_COMB.pep:*
    9:/cgn2_6/ptodata/1/paa/US085_COMB.pep:*
   10:/cgn2_6/ptodata/1/paa/US086_COMB.pep:*
   11:/cgn2_6/ptodata/1/paa/US087_COMB.pep:*
   12:/cgn2_6/ptodata/1/paa/US088_COMB.pep:*
   13:/cgn2_6/ptodata/1/paa/US089_COMB.pep:*
   14:/cgn2_6/ptodata/1/paa/US090_COMB.pep:*
   15:/cgn2_6/ptodata/1/paa/US091_COMB.pep:*
   16:/cgn2_6/ptodata/1/paa/US092_COMB.pep:*
   17:/cgn2_6/ptodata/1/paa/US093_COMB.pep:*
   18:/cgn2_6/ptodata/1/paa/US094_COMB.pep:*
   19:/cgn2_6/ptodata/1/paa/US095_COMB.pep:*
   20:/cgn2_6/ptodata/1/paa/US096_COMB.pep:*
   21:/cgn2_6/ptodata/1/paa/US097_COMB.pep:*
   22:/cgn2_6/ptodata/1/paa/US098_COMB.pep:*
   23:/cgn2_6/ptodata/1/paa/US099_COMB.pep:*
   24:/cgn2_6/ptodata/1/paa/US100_COMB.pep:*
   25:/cgn2_6/ptodata/1/paa/US101_COMB.pep:*
   26:/cgn2_6/ptodata/1/paa/US102_COMB.pep:*
   27:/cgn2_6/ptodata/1/paa/US60_COMB.pep:*

| No. | Score | % | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 7 | 1846 | 99.2 | 359 | 1 | PCT-US01-32619-1 | Sequence 1, Appli |
| 8 | 1846 | 99.2 | 359 | 23 | PCT-US01-32819-1 | Sequence 1, Appli |
| 9 | 1846 | 99.2 | 359 | 1 | US-09-899-292-1 | Sequence 1, Appli |
| 10 | 1846 | 99.2 | 359 | 1 | US-09-989-497-1 | Sequence 1, Appli |
| 11 | 1845 | 99.2 | 359 | 21 | US-09-791-537-72283 | Sequence 72283, A |
| 12 | 1844 | 99.1 | 359 | 18 | US-09-471-572-6 | Sequence 6, Appli |
| 13 | 1844 | 99.1 | 359 | 21 | US-09-791-537-95492 | Sequence 95492, A |
| 14 | 1842 | 99.0 | 359 | 37 | PCT-US01-32819-37 | Sequence 37, Appl |
| 15 | 1842 | 99.0 | 359 | 1 | PCT-US01-32619-27 | Sequence 27, Appl |
| 16 | 1842 | 99.0 | 359 | 37 | US-09-791-537-114678 | Sequence 114678, |
| 17 | 1842 | 99.0 | 359 | 23 | US-09-899-292-37 | Sequence 37, Appl |
| 18 | 1842 | 99.0 | 359 | 37 | US-09-989-497-37 | Sequence 37, Appl |
| 19 | 1837 | 98.8 | 359 | 1 | PCT-US01-32819-15 | Sequence 15, Appl |
| 20 | 1837 | 98.8 | 359 | 23 | US-09-899-292-15 | Sequence 15, Appl |
| 21 | 1837 | 98.8 | 359 | 23 | US-09-989-497-15 | Sequence 15, Appl |
| 22 | 1837 | 98.8 | 359 | 21 | US-09-791-537-78674 | Sequence 78674, A |
| 23 | 1830 | 98.4 | 359 | 21 | US-09-791-537-78011 | Sequence 78011, A |
| 24 | 1830 | 98.4 | 359 | 29 | US-09-679-664-29 | Sequence 29, Appl |
| 25 | 1827 | 98.2 | 359 | 33 | US-09-679-664-33 | Sequence 33, Appl |
| 26 | 1827 | 98.2 | 359 | 37 | US-09-679-664-37 | Sequence 37, Appl |
| 27 | 1827 | 98.2 | 359 | 20 | US-09-679-664-37 | Sequence 37, Appl |
| 28 | 1827 | 98.2 | 359 | 20 | US-09-679-664-47 | Sequence 47, Appl |
| 29 | 1422 | 98.2 | 359 | 20 | US-09-679-664-49 | Sequence 49, Appl |
| 30 | 1422 | 98.2 | 359 | 20 | US-09-679-664-50 | Sequence 50, Appl |
| 31 | 1418 | 98.2 | 359 | 20 | US-09-679-664-41 | Sequence 41, Appl |
| 32 | 1324 | 98.2 | 359 | 20 | US-09-679-664-45 | Sequence 45, Appl |
| 33 | 353 | 98.0 | 359 | 21 | US-09-791-537-120054 | Sequence 120054, |
| 34 | 1822 | 98.0 | 359 | 1 | PCT-US01-32619-9 | Sequence 9, Appli |
| 35 | 1822 | 98.0 | 359 | 1 | PCT-US01-32819-2 | Sequence 2, Appli |
| 36 | 1822 | 98.0 | 359 | 23 | US-09-899-292-9 | Sequence 9, Appli |
| 37 | 1822 | 98.0 | 359 | 23 | US-09-989-497-9 | Sequence 9, Appli |
| 38 | 1819 | 97.8 | 359 | 1 | PCT-US01-32819-2 | Sequence 2, Appli |
| 39 | 1819 | 97.8 | 359 | 1 | PCT-US01-32619-2 | Sequence 2, Appli |
| 40 | 1819 | 97.8 | 359 | 23 | US-09-989-497-2 | Sequence 2, Appli |
| 41 | 1819 | 97.8 | 359 | 23 | US-09-899-292-2 | Sequence 2, Appli |
| 42 | 1818 | 97.7 | 359 | 1 | US-09-984-292-2 | Sequence 2, Appli |
| 43 | 1818 | 97.7 | 359 | 1 | PCT-US01-32816-21 | Sequence 21, Appl |
| 44 | 1818 | 97.7 | 359 | 23 | US-09-984-292-21 | Sequence 21, Appl |
| 45 | 1818 | 97.7 | 359 | 23 | US-09-989-497-21 | Sequence 21, Appl |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1860 | 100.0 | 359 | 21 | US-09-791-537-57305 | Sequence 57305, A |
| 2 | 1860 | 100.0 | 359 | 22 | US-09-899-295-2 | Sequence 2, Appli |
| 3 | 1857 | 99.8 | 359 | 22 | US-09-471-572-8 | Sequence 8, Appli |
| 4 | 1857 | 99.8 | 359 | 21 | US-09-791-537-81282 | Sequence 81282, A |
| 5 | 1849 | 99.4 | 359 | 18 | US-09-471-572-7 | Sequence 7, Appli |
| 6 | 1849 | 99.4 | 359 | 21 | US-09-791-537-78690 | Sequence 78690, A |

RESULT 1
US-09-791-537-57305
Sequence 57305, Application US/09791537
GENERAL INFORMATION:
  APPLICANT: Bionomix, Inc.
  APPLICANT: Debe, Derek
  APPLICANT: Danzer, Joseph
  TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
  TITLE OF INVENTION: METHODS OF USE THEREOF
  FILE REFERENCE: 261/210
  CURRENT APPLICATION NUMBER: US/09/791,537
  CURRENT FILING DATE: 2001-02-22
  NUMBER OF SEQ ID NOS: 153055
  SOFTWARE: PatentIn version 3.0
  SEQ ID NO 57305
  LENGTH: 359
  TYPE: PRT
  ORGANISM: Homo sapiens
US-09-791-537-57305

ALIGNMENTS

Query Match 100.0%;  Score 1860;  DB 21;  Length 359;
Best Local Similarity 100.0%;  Pred. No. 1.1e-175;
Matches 359;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARELKILLLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARELKILLLGTGESGKSTFIKQMR 60

Qy    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

```
Db  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 2
US-09-899-295-2
Sequence 2, Application US/09899295
GENERAL INFORMATION:
  APPLICANT: Aventis Pharma Deutschland GmbH
  TITLE OF INVENTION: Process for identifying modulators of G protein coupled
  TITLE OF INVENTION: receptors
  FILE REFERENCE: AVE D-2000/A033 englisch
  CURRENT APPLICATION NUMBER: US/09/899,295
  CURRENT FILING DATE: 2001-07-06
  NUMBER OF SEQ ID NOS: 10
  SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
  LENGTH: 359
  TYPE: PRT
  ORGANISM: Mus musculus
US-09-899-295-2

  Query Match             100.0%;
  Best Local Similarity   100.0%;
  Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  Score 1860; DB 22; Pred. No. 1.1e-175; Length 359;

```
Qy   1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy 121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy 181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy 241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy 301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 3
US-09-471-572-8
Sequence 8, Application US/09471572
GENERAL INFORMATION:
  APPLICANT: Jones, Kenneth A.
  APPLICANT: Walker, Mary W.
  APPLICANT: Tamm, Joseph
  APPLICANT: Branchek, Theresa A.
  APPLICANT: Gerald, Christophe P.G.
  TITLE OF INVENTION: Chimeric G-Proteins And Uses Thereof
  FILE REFERENCE: 59896
  CURRENT APPLICATION NUMBER: US/09/471,572
  CURRENT FILING DATE: 1999-12-23
  NUMBER OF SEQ ID NOS: 45
  SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
  LENGTH: 359
  TYPE: PRT
  ORGANISM: Mus musculus
US-09-471-572-8

  Query Match             99.8%;
  Best Local Similarity   99.7%;
  Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
  Score 1857; DB 18; Pred. No. 2.1e-175; Length 359;

```
Qy   1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy 121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
       ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db 121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180

Qy 181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy 241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy 301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 4
US-09-791-537-81282
Sequence 81282, Application US/09791537
GENERAL INFORMATION:
  APPLICANT: Bionomix, Inc.
  APPLICANT: Debe, Derek
  APPLICANT: Danzer, Joseph
  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
  TITLE OF INVENTION: METHODS OF USE THEREOF
  FILE REFERENCE: 261/210
  CURRENT APPLICATION NUMBER: US/09/791,537
  CURRENT FILING DATE: 2001-02-22
  NUMBER OF SEQ ID NOS: 153055
  SOFTWARE: Patentin version 3.0
SEQ ID NO 81282
  LENGTH: 359
  TYPE: PRT
  ORGANISM: Mus musculus
US-09-791-537-81282

  Query Match             99.8%;
  Best Local Similarity   99.7%;
  Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
  Score 1857; DB 21; Pred. No. 2.1e-175; Length 359;

```
Qy   1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
```

RESULT 5
US-09-471-572-7
Sequence 7, Application US/09471572
GENERAL INFORMATION:
; APPLICANT: Jones, Kenneth A.
; APPLICANT: Walker, Mary W.
; APPLICANT: Tamm, Joseph
; APPLICANT: Branchek, Theresa A.
; APPLICANT: Gerald, Christophe P.G.
; TITLE OF INVENTION: Chimeric G-Proteins And Uses Thereof
; FILE REFERENCE: 59896
; CURRENT APPLICATION NUMBER: US/09/471,572
; CURRENT FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-471-572-7

Query Match             99.4%;  Score 1849;  DB 18;  Length 359;
Best Local Similarity   99.4%;  Pred. No. 1.3e-174;
Matches  357;  Conservative  1;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 6
US-09-791-537-78690
; Sequence 78690, Application US/0971537
GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 78690
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-791-537-78690

Query Match             99.4%;  Score 1849;  DB 21;  Length 359;
Best Local Similarity   99.4%;  Pred. No. 1.3e-174;
Matches  357;  Conservative  1;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 7
PCT-US01-32619-1
; Sequence 1, Application PC/TUS0132619
GENERAL INFORMATION:
; APPLICANT: SENOMYX, INC.
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-140309
; CURRENT APPLICATION NUMBER: PCT/US01/32619
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US01-32619-1

Query Match             99.4%;  Score 1846;  DB 1;  Length 359;
Best Local Similarity   99.4%;  Pred. No. 2.6e-174;
Matches  356;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLILLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 8
PCT-US01-32819-1
Sequence 1, Application PC/TUS0132819

GENERAL INFORMATION:
APPLICANT: YAO, YONG
APPLICANT: XU, HONG
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
FILE REFERENCE: 078003-0280649
CURRENT APPLICATION NUMBER: PCT/US01/32819
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: 60/243,770
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1
LENGTH: 359
TYPE: PRT
ORGANISM: Mus sp.

PCT-US01-32819-1

Query Match          99.28%;   Score 1846;  DB 1;   Length 359;
Best Local Similarity 99.28%;   Pred. No. 2.6e-174;
Matches 356;  Conservative 2;  Mismatches 1;  Indels 0;  Gaps 0;

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 9
US-09-984-292-1
Sequence 1, Application US/09984292

GENERAL INFORMATION:
APPLICANT: YAO, YONG
APPLICANT: XU, HONG
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
FILE REFERENCE: 078003-0280649
CURRENT APPLICATION NUMBER: US/09/984,292
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 60/243,770
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1
LENGTH: 359
TYPE: PRT
ORGANISM: Mus sp.

US-09-984-292-1

Query Match          99.28%;   Score 1846;  DB 23;   Length 359;
Best Local Similarity 99.28%;   Pred. No. 2.6e-174;
Matches 356;  Conservative 2;  Mismatches 1;  Indels 0;  Gaps 0;

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 10
US-09-989-497-1
Sequence 1, Application US/09989497

GENERAL INFORMATION:
APPLICANT: YAO, YONG
APPLICANT: XU, HONG
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
FILE REFERENCE: 078003-0280735
CURRENT APPLICATION NUMBER: US/09/989,497
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 09/984,292
PRIOR FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 60/243,770
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1
LENGTH: 359
TYPE: PRT

;
ORGANISM: Mus sp.
US-09-989-497-1

Query Match        99.2%;  Score 1846;  DB 23;  Length 359;
Best Local Similarity  99.2%;  Pred. No. 2.6e-174;
Matches 356;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 11
US-09-791-537-72283
Sequence 72283, Application US/0971537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 72283
LENGTH: 359
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-791-537-72283

Query Match        98.9%;  Score 1845;  DB 21;  Length 359;
Best Local Similarity  98.9%;  Pred. No. 3.3e-174;
Matches 355;  Conservative  3;  Mismatches  1;  Indels  0;  Gaps  0;

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMVRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

```
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 12
US-09-471-572-6
Sequence 6, Application US/09471572
GENERAL INFORMATION:
APPLICANT: Jones, Kenneth A.
APPLICANT: Walker, Mary W.
APPLICANT: Tamm, Joseph
APPLICANT: Branchek, Theresa A.
APPLICANT: Gerald, Christophe P.G.
TITLE OF INVENTION: Chimeric G-Proteins And Uses Thereof
FILE REFERENCE: 59896
CURRENT APPLICATION NUMBER: US/09/471,572
CURRENT FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 45
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
US-09-471-572-6

Query Match        99.1%;  Score 1844;  DB 18;  Length 359;
Best Local Similarity  99.2%;  Pred. No. 4.1e-174;
Matches 356;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV 359
```

RESULT 13
US-09-791-537-95492
Sequence 95492, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 95492

```
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-95492

  Query Match     Best Local Similarity   99.1%;   Score 1844;  DB 21;  Length 359;
  Matches 356;   Conservative  1;  Pred. No. 4.1e-174;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1 MTLESIMACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         ||||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEAARRINDEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 14
PCT-US01-32619-27
; Sequence 27, Application PC/TUS0132619
; GENERAL INFORMATION:
; APPLICANT: SENOMYX, INC.
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-140309
; CURRENT APPLICATION NUMBER: PCT/US01/32619
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-32619-27

  Query Match     Best Local Similarity   99.0%;   Score 1842;  DB 1;  Length 359;
  Matches 356;   Conservative  1;  Pred. No. 6.5e-174;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1 MTLESIMACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         ||||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEAARRINDEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 15
PCT-US01-32819-37
; Sequence 37, Application PC/TUS0132819
; GENERAL INFORMATION:
; APPLICANT: YAO, XU
; APPLICANT: XU, HONG
; APPLICANT: YAO, YONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: PCT/US01/32819
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-32819-37

  Query Match     Best Local Similarity   99.0%;   Score 1842;  DB 1;  Length 359;
  Matches 356;   Conservative  1;  Pred. No. 6.5e-174;  Mismatches 2;  Indels 0;  Gaps 0;

Qy    1 MTLESIMACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         ||||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEAARRINDEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKINYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||||||||||||||||||||| :||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

Search completed: June 30, 2003, 16:24:41
Job time : 319 secs

```
OM protein - protein search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:             June 30, 2003, 16:18:06 ; Search time 50 seconds
                                              (without alignments)
                                              1870.039 Million cell updates/sec

Title:              US-09-899-295-2
Perfect score:      1860
Sequence:           1 MTLESIMACCLSEEAKEARR..........VFAAVKDTILQLNLKEYNLV 359
Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           1197705 seqs, 260451061 residues

Total number of hits satisfying chosen parameters:    1197705

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Pending_Patents_AA_New:*
            1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
            2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
            3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
            4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
            5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
            6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
            7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1849 | 99.4 | 359 | 6 | US-10-405-027-3227 | Sequence 3227, Ap |
| 2 | 1849 | 99.4 | 359 | 7 | US-60-453-135-10419 | Sequence 10419, A |
| 3 | 1849 | 99.4 | 359 | 7 | US-60-453-050-10419 | Sequence 10419, A |
| 4 | 1849 | 99.4 | 359 | 7 | US-60-455-444-5597 | Sequence 5597, Ap |
| 5 | 1849 | 99.4 | 359 | 7 | US-60-465-241-5597 | Sequence 5597, Ap |
| 6 | 1849 | 99.4 | 359 | 7 | US-60-466-412-10419 | Sequence 10419, A |
| 7 | 1844 | 99.1 | 359 | 6 | US-10-258-561-2 | Sequence 2, Appl |
| 8 | 1817 | 97.7 | 353 | 1 | PCT-US03-02452-13 | Sequence 13, Appl |
| 9 | 1817 | 97.7 | 353 | 6 | US-10-352-843-13 | Sequence 13, Appl |
| 10 | 1687 | 90.7 | 387 | 5 | US-60-724-676A-49632 | Sequence 49632, A |
| 11 | 1687 | 90.7 | 387 | 5 | US-60-724-676-49632 | Sequence 49632, A |
| 12 | 1657 | 89.1 | 351 | 5 | US-60-949-016-7853 | Sequence 7853, Ap |
| 13 | 1535 | 82.5 | 355 | 6 | US-10-326-896-2 | Sequence 2, Appli |
| 14 | 1513 | 81.3 | 355 | 1 | PCT-US03-02452-12 | Sequence 12, Appl |
| 15 | 1513 | 81.3 | 355 | 6 | US-10-369-493-4982 | Sequence 4982, Ap |
| 16 | 1513 | 81.3 | 355 | 6 | US-10-352-843-12 | Sequence 12, Appl |
| 17 | 1104 | 59.4 | 252 | 5 | US-60-949-016-9408 | Sequence 9408, Ap |
| 18 | 1066 | 57.3 | 374 | 6 | US-10-468-002-4 | Sequence 4, Appl |
| 19 | 1066 | 57.3 | 374 | 6 | US-10-319-416-2 | Sequence 2, Appli |
| 20 | 1055 | 56.7 | 374 | 6 | US-10-299-442-20 | Sequence 20, Appl |
| 21 | 1055 | 56.7 | 374 | 6 | US-10-414-797-20 | Sequence 20, Appl |
| 22 | 1044 | 56.1 | 374 | 5 | US-60-468-002-2 | Sequence 2, Appli |
| 23 | 1044 | 56.1 | 374 | 5 | US-10-319-416-4 | Sequence 4, Appli |
| 24 | 1033 | 55.5 | 378 | 5 | US-60-949-016-7851 | Sequence 7851, Ap |
| 25 | 898.5 | 48.3 | 354 | 1 | PCT-US03-02452-14 | Sequence 14, Appl |
| 26 | 898.5 | 48.3 | 354 | 6 | US-10-352-843-14 | Sequence 14, Appl |
| 27 | 897.5 | 48.3 | 354 | 6 | US-10-180-930-2 | Sequence 2, Appli |
| 28 | 893.5 | 48.0 | 353 | 6 | US-10-219-051B-3379 | Sequence 3379, Ap |
| 29 | 893.5 | 48.0 | 353 | 6 | US-10-219-051B-3383 | Sequence 3383, Ap |
| 30 | 892.5 | 48.0 | 353 | 6 | US-10-219-051B-3377 | Sequence 3377, Ap |
| 31 | 892.5 | 48.0 | 353 | 6 | US-10-219-051B-3381 | Sequence 3381, Ap |
| 32 | 888 | 47.7 | 710 | 1 | PCT-US02-31059B-12 | Sequence 12, Appl |
| 33 | 881.5 | 47.4 | 354 | 5 | US-09-949-016-6727 | Sequence 6727, Ap |
| 34 | 881.5 | 47.4 | 354 | 5 | US-60-443-566-3898 | Sequence 3898, Ap |
| 35 | 881.5 | 47.4 | 354 | 5 | US-60-452-680-23965 | Sequence 23965, A |
| 36 | 881.5 | 47.4 | 354 | 7 | US-60-453-135-14917 | Sequence 14917, A |
| 37 | 881.5 | 47.4 | 354 | 7 | US-60-453-050-14917 | Sequence 14917, A |
| 38 | 881.5 | 47.4 | 354 | 7 | US-60-455-444-8098 | Sequence 8098, Ap |
| 39 | 881.5 | 47.4 | 354 | 7 | US-60-465-241-8098 | Sequence 8098, Ap |
| 40 | 881.5 | 47.4 | 354 | 7 | US-60-466-412-14917 | Sequence 14917, A |
| 41 | 877.5 | 47.2 | 353 | 6 | US-10-219-051B-11687 | Sequence 11687, A |
| 42 | 877 | 47.2 | 355 | 5 | US-09-949-016-9379 | Sequence 9379, Ap |
| 43 | 877 | 47.2 | 355 | 7 | US-60-453-135-9379 | Sequence 9379, Ap |
| 44 | 877 | 47.2 | 355 | 7 | US-60-453-050-9379 | Sequence 9379, Ap |
| 45 | 877 | 47.2 | 395 | 5 | US-60-466-412-9379 | Sequence 9379, Ap |
| . | 877 | 47.2 | 395 | 5 | US-09-949-016-11560 | Sequence 11560, A |

ALIGNMENTS

```
RESULT 1
US-10-405-027-3227
Sequence 3227, Application US/10405027
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS806P1
;  CURRENT APPLICATION NUMBER: US10/405,027
;  CURRENT FILING DATE: 2003-04-07
;  PRIOR APPLICATION NUMBER: 60/369,608
;  PRIOR FILING DATE: 2002-04-04
;  PRIOR APPLICATION NUMBER: 60/376,175
;  PRIOR FILING DATE: 2002-04-30
;  NUMBER OF SEQ ID NOS: 5810
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 3227
;  LENGTH: 359
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-405-027-3227

  Query Match             99.4%;  Score 1849;  DB 6;  Length 359;
  Best Local Similarity   99.4%;
  Matches  357;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy          1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
              ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db          1 MTLESIMACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy         61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy        121 VSAFENPYVDAIKSIWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTTQQDVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VSAFENPYVDAIKSIWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTTQQDVL 180

Qy        181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy        241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy        301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 2
US-60-453-135-10419
Sequence 10419, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: IAKOUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001456
CURRENT APPLICATION NUMBER: US/60/453,135
CURRENT FILING DATE: 2003-03-10
SEQ ID NO 10419
  LENGTH: 359
  TYPE: PRT
  ORGANISM: Homo sapiens
US-60-453-135-10419

Query Match    99.4%;    Score 1849;    DB 7;    Length 359;
Best Local Similarity    99.4%;    Pred. No. 4.1e-183;
Matches 357;    Conservative 1;    Mismatches 1;    Indels 0;    Gaps 0;

Qy        1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
            ||||||||||||||||||||||||||||||||  :|||||||||||||||||||||||||||
Db        1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy       61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy      121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy      181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy      241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy      301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 3
US-60-453-050-10419
Sequence 10419, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: LUKE, May
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001457
CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10
SEQ ID NO 10419
  LENGTH: 359
  TYPE: PRT
  ORGANISM: Homo sapiens
US-60-453-050-10419

Query Match    99.4%;    Score 1849;    DB 7;    Length 359;
Best Local Similarity    99.4%;    Pred. No. 4.1e-183;
Matches 357;    Conservative 1;    Mismatches 1;    Indels 0;    Gaps 0;

Qy        1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

RESULT 4
US-60-455-444-5597
Sequence 5597, Application US/60455444
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: BEGOVICH, Ann
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001455
CURRENT APPLICATION NUMBER: US/60/455,444
CURRENT FILING DATE: 2003-03-18
SOFTWARE: FastSEQ for Windows Version 4.0
NUMBER OF SEQ ID NOS: 50986
SEQ ID NO 5597
  LENGTH: 359
  TYPE: PRT
  ORGANISM: Homo sapiens
US-60-455-444-5597

Query Match    99.4%;    Score 1849;    DB 7;    Length 359;
Best Local Similarity    99.4%;    Pred. No. 4.1e-183;
Matches 357;    Conservative 1;    Mismatches 1;    Indels 0;    Gaps 0;

Qy        1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
            ||||||||||||||||||||||||||||||  :||||||||||||||||||||||||||||
Db        1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy       61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy      121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy      181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy      241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy      301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 5

Qy        1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

```
US-60-465-241-5597
  Sequence 5597, Application US/60465241
; GENERAL INFORMATION:
;   APPLICANT: CARGILL, Michele
;   APPLICANT: BEGOVICH, Ann
;   TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;   TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001468
;   CURRENT APPLICATION NUMBER: US/60/465,241
;   CURRENT FILING DATE: 2003-04-23
;   NUMBER OF SEQ ID NOS: 258418
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5597
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-60-465-241-5597

Query Match             99.4%;  Score 1849;  DB 7;  Length 359;
Best Local Similarity   99.4%;
Matches   357;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVHRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
           |||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db     1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy    61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLLDRVADPAYLPTQQDVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLLDRVADPAYLPTQQDVL 180

Qy   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359


RESULT 6
US-60-466-412-10419
  Sequence 10419, Application US/60466412
; GENERAL INFORMATION:
;   APPLICANT: CARGILL, Michele
;   APPLICANT: IAKOUBOVA, Olga
;   TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;   TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001466
;   CURRENT APPLICATION NUMBER: US/60/466,412
;   CURRENT FILING DATE: 2003-04-30
;   NUMBER OF SEQ ID NOS: 429241
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10419
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-60-466-412-10419

Query Match             99.4%;  Score 1849;  DB 7;  Length 359;
Best Local Similarity   99.4%;
Matches   357;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVHRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
           |||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db     1 MTLESIMACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
```

```
RESULT 7
US-10-258-561-2
  Sequence 2, Application US/10258561
; GENERAL INFORMATION:
;   APPLICANT: Young, Kathleen
;   APPLICANT: Howland, David S.
;   APPLICANT: Marquis, Karen L.
;   APPLICANT: Rosenzweig-Lipson, Sharon
;   APPLICANT: Cockett, Mark Ian
;   TITLE OF INVENTION: Transgenic Rat
;   FILE REFERENCE: 3655/1J046US2
;   CURRENT APPLICATION NUMBER: US/10/258,561
;   CURRENT FILING DATE: 2002-10-24
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: mouse G alpha q containing Gly to Ser substitution at codon 18
;   OTHER INFORMATION: as well as several other changes to include an EE epitope tag
US-10-258-561-2

Query Match             98.9%;  Score 1844;  DB 6;  Length 359;
Best Local Similarity   98.9%;
Matches   355;  Conservative   2;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVHRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
           |||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||
Db     1 MTLESIMACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy    61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLLDRVADPAYLPTQQDVL 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLLDRVADPAYMPTEODVL 180

Qy   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

```
Query Match             99.8%;  Score 1849;  DB 7;  Length 359;
Best Local Similarity   99.4%;  Pred. No. 4.1e-183;
Matches   357;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLLDRVADPAYLPTQQDVL 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLLDRVADPAYLPTQQDVL 180

Qy   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

```
Query Match             99.4%;  Score 1849;  DB 7;  Length 359;
Best Local Similarity   99.4%;  Pred. No. 4.1e-183;
Matches   357;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
           |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db     1 MTLESIMACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
```

Db        301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 8
PCT-US03-02452-13
; Sequence 13, Application PCT/US0302452
; GENERAL INFORMATION:
; APPLICANT: Moore, Lisa
; APPLICANT: Kindt, Rachel
; APPLICANT: Kopczynski, Jenny
; APPLICANT: Doberstein, Stephen
; APPLICANT: Cockett, Mark
; APPLICANT: Ramanathan, Chandra
; APPLICANT: Lodge, Nicholas
; APPLICANT: Fitzgerald, Kevin
; APPLICANT: Stouch, Terry
; TITLE OF INVENTION: MOLECULES THAT MODULATE G(ALPHA)q ACTIVITY AND METHODS OF
; TITLE OF INVENTION: TREATING URINARY INCONTINENCE
; FILE REFERENCE: 5624-277-228
; CURRENT APPLICATION NUMBER: PCT/US03/02452
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/352720
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: G-protein of the invention
PCT-US03-02452-13

Query Match       97.7%;    Score 1817;   DB 1;  Length 353;
Best Local Similarity    99.2%;   Pred. No. 8.4e-180;
Matches  350;  Conservative    1;  Mismatches   2;  Indels    0;  Gaps    0;

Qy      7  MACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG  66
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG  60

Qy     67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  120

Qy    127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  186
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  180

Qy    187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy    247  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  300

Qy    307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV  353

RESULT 9
US-10-352-843-13
; Sequence 13, Application US/1035Z843
; GENERAL INFORMATION:
; APPLICANT: Moore, Lisa
; APPLICANT: Kindt, Rachel
; APPLICANT: Kopczynski, Jenny
; APPLICANT: Doberstein, Stephen
; APPLICANT: Cockett, Mark
; APPLICANT: Ramanathan, Chandra
; APPLICANT: Lodge, Nicholas
; APPLICANT: Fitzgerald, Kevin
; APPLICANT: Stouch, Terry
; TITLE OF INVENTION: MOLECULES THAT MODULATE G(ALPHA)q ACTIVITY AND METHODS OF
; TITLE OF INVENTION: TREATING URINARY INCONTINENCE
; FILE REFERENCE: 5624-277-999
; CURRENT APPLICATION NUMBER: US/10,352,843
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: US 60/352720
; PRIOR FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: G-protein of the invention
US-10-352-843-13

Query Match       97.7%;    Score 1817;   DB 6;  Length 353;
Best Local Similarity    99.2%;   Pred. No. 8.4e-180;
Matches  350;  Conservative    1;  Mismatches   2;  Indels    0;  Gaps    0;

Qy'     7  MACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG  66
           |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db      1  MACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG  60

Qy     67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  120

Qy    127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  186
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  180

Qy    187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy    247  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  300

Qy    307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV  353

RESULT 10
US-09-724-676-49632
; Sequence 49632, Application US/0972Z676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09,724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49632
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-49632

Query Match       90.7%;    Score 1687;   DB 5;  Length 387;
Best Local Similarity    89.4%;   Pred. No. 3e-166;
Matches  321;  Conservative   23;  Mismatches  15;  Indels    0;  Gaps    0;

Qy      1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR  60
           |||||:|:|||||:|||||||||||| ||||: :|||||||||||||||||||||||||||
Db      1  MTLESMMACCLSDEVKESKRINAEIEKQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR  60

RESULT 11
US-09-724-676A-49632
Sequence 49632, Application US/0972467

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 12918I.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 49632
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-49632

Query Match             90.7%;  Score 1687; DB 5; Length 387;
Best Local Similarity   89.4%;  Pred. No. 3e-166;
Matches 321; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR   60
        |||||||:::||:|| ::|||::||||||||||||||||||||||||||||||||||||
Db    1 MTLESMMACCLSDEVKESKRINAEIEKQLRDKRDARRELKLILLGTGESGKSTFIKQMR   60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
        |||:|:|||||||||||||||||||||||||||:|||||||||||||:||||||||||||
Db   61 IIHGAGYSEEDKRGFTKLVYQNIFTAMQAMIRAMETLKILYKYEQNKANALLIREVDVEK  120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVL  180
        :||:||| |:||||||||:|:|| |||||||||||||||||||| |:| | |||||||||
Db  121 VTTFEHQYVSAIKTLWEDPGIQECYDRRREYQLSDSAKYILTDVDRIATLGYLPTQQDVL  180

Qy  181 RVRVPFTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
        |||||||||||||||||:|::|||||||||||||||||||||||:|||||||||||||||
Db  181 RVRVPFTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240

Qy  241 ESDNENRMEESKALFRTIIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300
        ||||||||||||||||||||||||||||||||||||||||::|:|||||||||||||||||
Db  241 ESDNENRMEESKALFRTIIITYPWFQNSSVILFLNKKDLLEDKILYSHLVDYFPEFDGPQR  300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLL  359
        ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||:
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLL  359
```

RESULT 12
US-09-949-016-7853
Sequence 7853, Application US/09949016

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7853
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7853

Query Match             89.1%;  Score 1657; DB 5; Length 351;
Best Local Similarity   89.7%;  Pred. No. 3.3e-163;
Matches 315; Conservative 21; Mismatches 15; Indels 0; Gaps 0;

```
Qy    9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMRIIHGSGYS   68
        |||:||:|| ::|||::|||||||||||||||||||||||||||||||||||:|:|||||
Db    1 CCLSDEVKESKRINAEIEKQLRDKRDARRELKLILLGTGESGKSTFIKQMRIIHGAGYS   60

Qy   69 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPY  128
        |||||||||||||||||||||||||:|||||||||||||:|||||||||||||||:|||| 
Db   61 EEDKRGFTKLVYQNIFTAMQAMIRAMETLKILYKYEQNKANALLIREVDVEKVTTFEHQY  120

Qy  129 VDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVLRVRVPFTTG  188
        :|||| :|:|||||| ||||||||||||:|||| | |:| | |||||||||||||||||||
Db  121 VSAIKTLWEDPGIQECYDRRREYQLSDSAKYILTDVDRIATLGYLPTQQDVLRVRVPFTTG  180

Qy  189 IIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM  248
        ||||||||:|::||||||||||||||||||||||:|||||||||||||||||||||||||
Db  181 IIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM  240

Qy  249 EESKALFRTIIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREF  308
        ||||||||||||||||||||||||||||||||::|:||||||||||||||||||||||||
Db  241 EESKALFRTIIITYPWFQNSSVILFLNKKDLLEDKILYSHLVDYFPEFDGPQRDAQAAREF  300

Qy  309 ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  351
```

RESULT 13
US-10-326-896-2
Sequence 2, Application US/10326896

; GENERAL INFORMATION:
; APPLICANT: Keith D. Allen
; APPLICANT: Qin Zhang
; TITLE OF INVENTION: GNA14 G-Protein Signaling Protein
; TITLE OF INVENTION: Subunit Gene Disruptions, Compositions and Methods Related
; TITLE OF INVENTION: Thereto
; FILE REFERENCE: R-696
; CURRENT APPLICATION NUMBER: US/10/326,896
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-326-896-2

Query Match             82.5%;  Score 1535; DB 6; Length 355;
Best Local Similarity   82.1%;  Pred. No. 1.5e-150;
Matches 288; Conservative 35; Mismatches 28; Indels 0; Gaps 0;

```
Qy    9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMRIIHGSGYS   68
        |||| | ||::|:|| :||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 14
Sequence 12, Application PC/TUS030302452
GENERAL INFORMATION:
; APPLICANT: Moore, Lisa
; APPLICANT: Kindt, Rachel
; APPLICANT: Kopczynski, Jenny
; APPLICANT: Doberstein, Stephen
; APPLICANT: Cockett, Mark
; APPLICANT: Ramanathan, Chandra
; APPLICANT: Lodge, Nicholas
; APPLICANT: Fitzgerald, Kevin
; APPLICANT: Stouch, Terry
; TITLE OF INVENTION: MOLECULES THAT MODULATE G(ALPHA)q ACTIVITY AND METHODS OF
; TITLE OF INVENTION: TREATING URINARY INCONTINENCE
; FILE REFERENCE: 5624-277-228
; CURRENT APPLICATION NUMBER: PCT/US03/02452
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/352720
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.2
;
; SEQ ID NO 12
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
PCT-US03-02452-12

Query Match          81.3%;   Score 1513;  DB 1;  Length 355;
Best Local Similarity   82.0%;   Pred. No. 3e-148;
Matches  291;  Conservative   26;  Mismatches   36;  Indels    2;  Gaps    1;

Qy      5 CCLSAEEKESQRISAEIERHVRDKKDARRELKLILLLGTGESGKSTFIKQMRIIHGSGYS 64
          ||| |:: | ||| |::| | |||| |||:|||||||||||||||||||||||||||||
Db    241 ENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQA 300

Qy     69 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPY 128
          || || ||||||||||||||||||: |||||  ||| |||||||||||||||: :||
Db    305 AREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

Qy    129 VDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVPTTG 188
          || | ||||||| ||||||||||||||||||||||||||||||||||||| |:||
Db     65 DEDRKGFTKLVQONIFTAMQAMIRAMDTLRIQYMCEQNENAQIIREVEVDKVFALSRDQ 124

Qy    189 IIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
          |||:|| | ||||||||||||||||||||||||||||:||  ||:||||||||||||||
Db    185 IIEYPFDLENIIFRMVDVGGQQRSERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244

Qy    249 EESKALFRTIITYPWFQNSSVILFLNKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAAREF 308
          |||||||||||||||||| ::|||||||:|||||||||||||||||||||||:||:| |
Db    245 EESKALFRTIITYPWFLNSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAAARDF 304

Qy    309 ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV 359
          ||||:| | ||| |||| |||:||||||||||||||||||:|:|||:|||
Db    305 ILKLYQDDONPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

RESULT 15
US-10-369-493-4982
Sequence 4982, Application US/10369493
GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-28
;
; SEQ ID NO 4982
; NUMBER OF SEQ ID NOS: 47374
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-4982

Query Match          81.3%;   Score 1513;  DB 6;  Length 355;
Best Local Similarity   82.0%;   Pred. No. 3e-148;
Matches  291;  Conservative   26;  Mismatches   36;  Indels    2;  Gaps    1;

Qy      7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLLGTGESGKSTFIKQMRIIHGSG 66
          |||||||||| ||||: ||||||| ||| |||||||||||||||||||||||||||||||
Db      1 MACCLSEEAREQKRINQEIEKQLQRDKRNARRELKLILLLGTGESGKSTFIKQMRIIHGQG 60

Qy     67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPY--KYEHNKAHAQLVREVDVEKVSAF 124
          |:|||||||| ||||||||||||||||||||||||  |||| |||||||||||||||||
Db     61 YSEEDKRAHIRLVYQNVFMAIQSMIRAMDTLDIKFGNESEELQEKAAVVREVDFESVTSF 120

Qy    125 ENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRV 184
          |||| | ||||||||||| ||||||||||||||||||||||||||||||||||||| |:|
Db    121 EEPYVSYIKELWEDSGIQECYDRRREYQLTDSAKYYLSDLRRLAVPDYLPTEQDILRVRV 180

Qy    185 PTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDN 244
          |||||||:|| | ||||||||||||||||||||||||||||:|| ||:|||||||||||
Db    181 PTTGIIEYPFDLEQIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVECDN 240

Qy    245 ENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPPRDPIA 304
          :||||||||||||||||||| ::|||||||||||||||||||:| |||||||||:||:|
Db    241 ENRMEESKALFRTIITYPWFTNSSVILFLNKKDLLEEKILYSHLADYFPEYDGPPRDPIA 300

Qy    305 AREFILKMFVDLNPDADKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          |||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db    301 AREFILKMFVDLNPDADKIIYSHFTCATDTENIRFVFAAVKDTILQHNLKEYNLV 355

Search completed: June 30, 2003, 16:25:31
Job time : 52 secs

OM protein - protein search, using sw model

Run on:             June 30, 2003, 16:09:36 ; Search time 39 Seconds
                                              (without alignments)
                                              1226.589 Million cell updates/sec

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2003 Compugen Ltd.

Title:             US-09-899-295-2
Perfect score:     1860
Sequence:          1 MTLESIMACCLSEBAKEARR..........VFAAVKDTILQLNLKEYNLV 359

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:        908470

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :         A_Geneseq_101002:*
         1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
         2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
         3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
         4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
         5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
         6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
         7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
         8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
         9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
        10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
        11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
        12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
        13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
        14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
        15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
        16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
        17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
        18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
        19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
        20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
        21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
        22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
        23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

SUMMARIES

| Result No. | Score | % Query Match | Length DB | ID | Description |
|---|---|---|---|---|---|
| 1 | 1860 | 100.0 | 359 | 23 | AAM48816 | Murine G-protein 6 |
| 2 | 1849 | 99.4 | 359 | 23 | ABG68610 | Human G protein al |
| 3 | 1846 | 99.2 | 359 | 23 | ABG68584 | Mouse G protein al |
| 4 | 1844 | 99.1 | 359 | 23 | ABG68598 | Human G protein al |
| 5 | 1827 | 98.2 | 1276 | 20 | AAY49127 | phCaR/hmGluR2-CaR |
| 6 | 1827 | 98.2 | 1276 | 23 | AAO15093 | Human phCaR/hmGluR |
| 7 | 1827 | 98.2 | 1394 | 20 | AAY49129 | pmGluR2/CaR*Galpha |
| 8 | 1827 | 98.2 | 1394 | 23 | AAO15095 | Human pmGluR2-CaR* |
| 9 | 1827 | 98.2 | 1397 | 20 | AAY49134 | pmGluR2/CaR*Galpha |
| 10 | 1827 | 98.2 | 1397 | 23 | AAO15100 | Human pmGluR2-CaR* |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| | 11 | 1827 | 98.2 | 1402 | 23 | AAO15105 | Human ph2SPMGluR3- |
| | 12 | 1827 | 98.2 | 1418 | 20 | AAY49131 | mGluR8/CaR*GalphaQ |
| | 13 | 1827 | 98.2 | 1421 | 23 | AAO15097 | Human mGluR8-CaR*G |
| | 14 | 1827 | 98.2 | 1422 | 20 | AAO15103 | Human phmGluR8-CaR |
| | 15 | 1827 | 98.2 | 1323 | 20 | AAO15102 | Human phmGluR4-CaR |
| | 16 | 1823 | 98.0 | 1323 | 20 | AAY49133 | GABA-BR1a*Gqo5 fus |
| | 17 | 1823 | 98.0 | 1323 | 23 | AAO15099 | Human GABA-BR1a*Gq |
| | 18 | 1822 | 98.0 | 359 | 23 | ABG68592 | Mouse GABA-BR2*Gq |
| | 19 | 1819 | 97.8 | 353 | 23 | ABG68585 | Mouse G protein al |
| | 20 | 1818 | 97.7 | 359 | 23 | AAY49125 | Chimeric Gq15 prot |
| | 21 | 1817 | 97.7 | 353 | 22 | ABG68604 | G protein-coupled |
| | 22 | 1817 | 97.6 | 1303 | 20 | AAH99072 | Human G-protein al |
| | 23 | 1816 | 97.6 | 1303 | 23 | ABB09281 | Human G-protein al |
| | 24 | 1816 | 97.6 | 353 | 23 | AAY49132 | GABA-BR2*Gq05 fusi |
| | 25 | 1815 | 97.6 | 359 | 23 | AAO15098 | Human GABA-BR2*Gqo |
| | 26 | 1814 | 97.5 | 359 | 23 | ABG68594 | Mouse G protein al |
| | 27 | 1814 | 97.5 | 353 | 23 | ABG68599 | Human G protein al |
| | 28 | 1814 | 97.5 | 359 | 23 | ABG68609 | Murine G-protein 6 |
| | 29 | 1813 | 97.5 | 359 | 23 | AAM48818 | Murine G-protein 6 |
| | 30 | 1811 | 97.4 | 359 | 23 | ABG68608 | Human G protein al |
| | 31 | 1811 | 97.4 | 359 | 23 | ABG68586 | Mouse G protein al |
| | 32 | 1810 | 97.3 | 353 | 23 | AAM48817 | Murine G-protein 6 |
| | 33 | 1807 | 97.2 | 359 | 23 | ABG68603 | Human G protein al |
| | 34 | 1806 | 97.1 | 359 | 23 | ABG68609 | Human G protein al |
| | 35 | 1804 | 97.0 | 359 | 23 | ABG68593 | Mouse G protein al |
| | 36 | 1803 | 96.9 | 359 | 23 | ABG68605 | Mouse G protein al |
| | 37 | 1800 | 96.8 | 359 | 23 | ABG68606 | Human G protein al |
| | 38 | 1796 | 96.6 | 359 | 23 | ABG68595 | Human G protein al |
| | 39 | 1795 | 96.5 | 353 | 23 | ABG68591 | Mouse G protein al |
| | 40 | 1793 | 96.4 | 359 | 23 | ABG68596 | Mouse G protein al |
| | 41 | 1791 | 96.3 | 353 | 23 | ABG68600 | Mouse G protein al |
| | 42 | 1788 | 96.1 | 353 | 23 | ABG68588 | Mouse G protein al |
| | 43 | 1784 | 95.9 | 353 | 23 | ABG68587 | Mouse G protein al |
| | 44 | 1779 | 95.6 | 353 | 23 | ABG68607 | Human G protein al |
| | 45 | 1776 | 95.5 | 353 | 23 | ABG68597 | Mouse G protein al |

ALIGNMENTS

RESULT 1
AAM48816
ID    AAM48816 standard; Protein; 359 AA.
XX
AC    AAM48816;
XX
DT    24-APR-2002  (first entry)
XX
DE    Murine G-protein 6qi4myr.
XX
KW    Mouse; G-protein coupled receptor modulator; G-protein 6qi4myr;
KW    signal transduction.
XX
OS    Mus musculus.
XX
PN    WO200204665-A2.
XX
PD    17-JAN-2002.
XX
PF    05-JUL-2001; 2001WO-EP07667.
XX
PR    08-JUL-2000; 2000DE-1033353.
XX
PA    (AVET ) AVENTIS PHARMA DEUT GMBH.
XX
PI    Kostenis E;
XX
DR    WPI; 2002-148182/19.
DR    N-PSDB; ABA97518.
XX
PT    Identifying compounds that modify activity of signal transduction
PT    pathways, useful potentially as therapeutic agents, by screening with
PT    cells that contain hybrid G proteins

## Left column — RESULT 2

```
RESULT 2
ID   ABG68610      standard; Protein; 359 AA.
AC   ABG68610;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human G protein alpha sub-unit q family #2.
XX
KW   G_q protein; sensory signaling; chemoreceptor; tastant; olfactant;
KW   pheromone; G protein alpha sub-unit; q family; G alpha q.
XX
OS   Homo sapiens.
XX
XX
PN   WO200236622-A2.
XX
PD   10-MAY-2002.
XX
PF   30-OCT-2000; 2000US-243770P.
XX
PR   24-OCT-2001; 2001WO-US32619.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Yao Y,   Xu H;
XX
DR   WPI; 2002-519234/55.
XX
PT   New G-alpha-q protein variants, useful for analyzing and discovering
     agonists or antagonists of chemoreceptors, such as G protein coupled
     receptors involved in sensing of tastants, olfactants or pheromones
XX
PS   Claim 22; Page 24-25; 34pp; German.
XX
CC   The present invention relates to a method of identifying compounds which
CC   are capable of modifying the activity of a signal transduction pathway
CC   which is dependent upon a G-protein coupled receptor. These compounds may
CC   include the novel G-proteins 6q14myr (shown here), 6q15myr, 6q14, 6qs5
CC   and Galpha16.
XX
SQ   Sequence  359 AA;

Query Match     100.0%;  Score 1860; DB 23;  Length 359;
Best Local Similarity  100.0%;  Pred. No. 3.6e-176;
Matches 359; Conservative  0; Mismatches  0; Indels  0; Gaps  0;

Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVL 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVL 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

## Right column — RESULT 3

```
RESULT 3
ID   ABG68584      standard; Protein; 359 AA.
AC   ABG68584;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Mouse G protein alpha sub-unit q family.
XX
KW   G_q protein; sensory signaling; chemoreceptor; tastant; olfactant;
KW   pheromone; G protein alpha sub-unit; q family; G alpha q.
XX
OS   Mus musculus.
XX
XX
PN   WO200236622-A2.
XX
PD   10-MAY-2002.
XX
PF   30-OCT-2000; 2000US-243770P.
XX
PR   24-OCT-2001; 2001WO-US32619.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Yao Y,   Xu H;
XX
DR   WPI; 2002-519234/55.
XX
PS   Disclosure; Fig 2; 32pp; English.
XX
CC   The invention describes an isolated variant of a G_q protein, which
CC   exhibits increased promiscuity relative to the corresponding G_q protein.
CC   The variant is used to identify a compound that modulates sensory
CC   signaling in sensory cells and to identify a compound that interacts
CC   with the G_q variant protein. The G_q protein variant is useful for
CC   analysing and discovering agonists or antagonists of chemoreceptors, such
CC   as G protein coupled receptors involved in sensing of tastants,
CC   olfactants or pheromones. This is the amino acid sequence of a G protein
CC   alpha sub-unit q family (G alpha q) protein that can functionally couple
CC   to sensory cell receptors such as taste GPCR's (G protein-coupled
CC   receptors) and olfactory GPCR's in an overly promiscuous manner.
XX
SQ   Sequence  359 AA;

Query Match     99.4%;  Score 1849; DB 23;  Length 359;
Best Local Similarity  99.4%;  Pred. No. 4.4e-175;
Matches 357; Conservative  1; Mismatches  1; Indels  0; Gaps  0;

Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
        ||||||||||||||||||||||||||| ::||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVL 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVL 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

## RESULT 4

```
RESULT 4
ABG68598
ID   ABG68598 standard; Protein; 359 AA.
XX
AC   ABG68598;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human G protein alpha sub-unit q family #1.
XX
KW   G_q protein; sensory signaling; chemoreceptor; tastant; olfactant;
KW   pheromone; G protein alpha sub-unit q family; G alpha q.
XX
OS   Homo sapiens.
XX
PN   WO200236622-A2.
XX
PD   10-MAY-2002.
XX
PF   24-OCT-2001; 2001WO-US32619.
XX
PR   30-OCT-2000; 2000US-243770P.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Yao Y, Xu H;
XX
DR   WPI; 2002-519234/55.
XX
PT   New G-alpha-q protein variants, useful for analyzing and discovering
PT   agonists or antagonists of chemoreceptors, such as G protein coupled
PT   receptors involved in sensing of tastants, olfactants or pheromones -
XX
PS   Claim 13; Fig 1; 32pp; English.
XX
CC   The invention describes an isolated variant of a G_q protein, which
CC   exhibits increased promiscuity relative to the corresponding G_q protein.
CC   The variant is used to identify a compound that modulates sensory
CC   signaling in sensory cells and to identify a compound that interacts
CC   with the G_q variant protein. The G_q protein variant is useful for
CC   analysing and discovering agonists or antagonists of chemoreceptors, such
CC   as G protein coupled receptors involved in sensing of tastants, such
CC   olfactants or pheromones. This is the amino acid sequence of a G protein
CC   alpha sub-unit q family (G alpha q) protein that can functionally couple
CC   to sensory cell receptors such as taste GPCR's (G protein-coupled
CC   receptors) and olfactory GPCR's in an overly promiscuous manner.
XX
SQ   Sequence   359 AA;

Query Match             99.2%;  Score 1846;  DB 23;  Length 359;
Best Local Similarity   99.2%;  Pred. No. 8.8e-175;
Matches  356;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         |||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

## RESULT 5

```
RESULT 5
AAY49127
ID   AAY49127 standard; Protein; 1276 AA.
XX
AC   AAY49127;
XX
DT   07-JAN-2000  (first entry)
XX
DE   phCaR/hmGluR2*Gq15 fusion construct protein sequence.
XX
KW   G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
KW   metabotropic glutamate receptor; GABABR; GABA receptor; stroke;
KW   gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW   spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW   Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW   cognitive disorder.
XX
OS   Homo sapiens.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Yao Y, Xu H;
XX
DR   WPI; 2002-519234/55.
XX
PT   New G-alpha-q protein variants, useful for analyzing and discovering
PT   agonists or antagonists of chemoreceptors, such as G protein coupled
PT   receptors involved in sensing of tastants, olfactants or pheromones -
XX
PS   Claim 13; Fig 1; 32pp; English.
XX
CC   The invention describes an isolated variant of a G_q protein, which
CC   exhibits increased promiscuity relative to the corresponding G_q protein.
CC   The variant is used to identify a compound that modulates sensory
CC   signaling in sensory cells and to identify a compound that interacts
CC   with the G_q variant protein. The G_q protein variant is useful for
CC   analysing and discovering agonists or antagonists of chemoreceptors, such
CC   as G protein coupled receptors involved in sensing of tastants, such
CC   olfactants or pheromones. This is the amino acid sequence of a G protein
CC   alpha sub-unit q family (G alpha q) protein that can functionally couple
CC   to sensory cell receptors such as taste GPCR's (G protein-coupled
CC   receptors) and olfactory GPCR's in an overly promiscuous manner.
XX
SQ   Sequence   359 AA;

Query Match             99.1%;  Score 1844;  DB 23;  Length 359;
Best Local Similarity   99.2%;  Pred. No. 1.4e-174;
Matches  356;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         |||||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV 359
```

XX
PN   WO9951641-A1.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;  99WO-US07333.
XX
PR   03-APR-1998;  98US-0080671.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI   Simin RT;
XX
DR   WPI; 1999-610995/52.
DR   N-PSDB; AAZ31058.
XX
PS   Example 1: Fig 12; 255pp; English.
XX
PT   New G-protein fusion receptors and chimeras containing domains from
PT   different receptors, used to screen for modulators, potentially useful
PT   e.g. for treating or preventing stroke or Alzheimer's disease -
XX
CC   The invention relates to G-protein fusion receptors (I) comprising:
CC   (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC   intracellular (ICD) domains, each chosen independently from a CaR
CC   (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
CC   (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC   the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC   linker. (1), and recombinant chimeric receptors (CR) without the GP
CC   component, are used to assess function of the various domains and to
CC   identify compounds (e.g. allosteric modulators or antagonists) that act
CC   on these domains. The modulators are potentially useful for treating or
CC   preventing diseases associated with the receptors, e.g. stroke, head or
CC   spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC   Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC   disorders and depression. Nucleic acid (II) that encodes (I) is used:
CC   (1) for recombinant production of corresponding proteins; and (2) to
CC   produce cells used in screening for modulators. Use of CaR and mGluR
CC   domains allows presentation of GABABR domains, to a binding agent, in a
CC   form more like the natural domain structure compared with use of
CC   incomplete receptors, lacking one or more domains. By shuffling different
CC   domains, agents can be identified that affect particular domains of a
CC   receptor.
XX
SQ   Sequence  1276 AA;

Query Match             98.2%;  Score 1827;  DB 20;  Length 1276;
Best Local Similarity   98.6%;  Pred. No. 4.3e-172;
Matches 353;  Conservative 2;  Mismatches 3;  Indels 0;  Gaps 0;

```
Qy      1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR   60
           |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db    918  MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLILLGTGESGKSTFIKQMR  977

Qy     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 1037

Qy    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1097

Qy    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 1157

Qy    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1217

Qy    301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL  358
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db   1218  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 1275
```

Db   1218  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL 1275

RESULT 6
AA015093
ID   AA015093 standard; Protein; 1276 AA.
XX
AC   AA015093;
XX
DT   22-AUG-2002  (first entry)
XX
OS   Chimeric - Homo sapiens.
XX
PN   WO2002229033-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001;  2001WO-US31074.
XX
PR   03-OCT-2000;  2000US-0679664.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI   Simin RT;
XX
DR   WPI; 2002-330170/36.
XX
PS   Example 3: Fig 12; 168pp; English.
XX
PT   Novel G-protein fusion receptor, useful for identifying modulators of
PT   CaR, mGluR and GABAB, comprises G-protein joined to the intracellular
PT   domain of the receptor -
XX
CC   The invention comprises G-protein fusion receptors - comprising
CC   extracellular, transmembrane and intracellular domains similar to CaR,
CC   mGluR or GABAB receptor sequences. The G-protein fusion receptors of the
CC   invention may also possess a linker joined to the carboxy terminus of the
CC   intracellular domain, and a G-protein joined to the linker. The G-protein
CC   fusion receptors of the invention are useful for identifying modulators
CC   of CaR, mGluR and GABAB for use in treating associated conditions. The
CC   present amino acid sequence was used in the production of the invention.
XX
SQ   Sequence  1276 AA;

Query Match             98.2%;  Score 1827;  DB 23;  Length 1276;
Best Local Similarity   98.6%;  Pred. No. 4.3e-172;
Matches 353;  Conservative 2;  Mismatches 3;  Indels 0;  Gaps 0;

```
Qy      1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR   60
           |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db    918  MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLILLGTGESGKSTFIKQMR  977

Qy     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 1037

Qy    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1097

Qy    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 1157

Qy    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1217
```

```
Db   1158 ESDNENREMEESKALFRTITTYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1217
Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
Db   1218 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL 1275
```

```
RESULT 7
AAY49129
ID   AAY49129 standard; Protein; 1394 AA.
XX
AC   AAY49129;
XX
DT   07-JAN-2000   (first entry)
XX
DE   pmGluR2/CaR*Galphaqi5 fusion construct protein sequence.
XX
KW   G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
KW   metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW   gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW   spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW   Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW   cognitive disorder.
XX
OS   Homo sapiens.
XX
PN   WO9951641-A1.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999; 99WO-US07333.
XX
PR   03-APR-1998; 98US-0080671.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI   Simin RT;
XX
DR   WPI; 1999-610995/52.
DR   N-PSDB; AA231060.
XX
PT   New G-protein fusion receptors and chimeras containing domains from
PT   different receptors, used to screen for modulators, potentially useful
PT   e.g. for treating or preventing stroke or Alzheimer's disease -
XX
PS   Example 1; Fig 12; 255pp; English.
XX
CC   The invention relates to G-protein fusion receptors (I) comprising:
CC   (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC   intracellular (ICD) domains, each chosen independently from a CaR
CC   (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
CC   (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC   the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC   linker. (1), and recombinant chimeric receptors (CR) without the GP
CC   component, are used to assess function of the various domains and to
CC   identify compounds (e.g. allosteric modulators or antagonists) that act
CC   on these domains. The modulators are potentially useful for treating or
CC   preventing diseases associated with the receptors, e.g. stroke, head or
CC   spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC   Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC   disorders and depression. Nucleic acid (II) that encodes (I) is used:
CC   (1) for recombinant production of corresponding proteins; and (2) to
CC   produce cells used in screening for modulators. Use of CaR and mGluR
CC   domains allows presentation of GABABR domains, to a binding agent, in a
CC   form more like the natural domain structure compared with use of
CC   incomplete receptors, lacking one or more domains. By shuffling different
CC   domains, agents can be identified that affect particular domains of a
CC   receptor.
XX
SQ   Sequence 1394 AA;
```

Query Match 98.2%; Score 1827; DB 20; Length 1394;

```
Best Local Similarity 98.6%;
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy     1 MTLESIMACCLSEAKEARRINDEIEHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60
         |||||||||:|||||||||||||| :||||||||||||||||||||||||||||||||
Db  1036 MTLESIMACCLSEAKEARRINDEIERQLRRDKRDARRELKLILLGTGESGKSTFIKQMR 1095

Qy    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1096 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 1155

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1156 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1215

Qy   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1216 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 1275

Qy   241 ESDNENRMEESKALFRTITTYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1276 ESDNENRMEESKALFRTITTYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1335

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db  1336 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL 1393
```

```
RESULT 8
AAO15095
ID   AAO15095 standard; Protein; 1394 AA.
XX
AC   AAO15095;
XX
DT   22-AUG-2002   (first entry)
XX
DE   Human pmGluR2-CaR*G-alpha-qi5 fusion construct protein.
XX
KW   Human; G-protein fusion receptor; extracellular domain;
KW   transmembrane domain; intracellular domain; CaR; mGluR; GABABR;
KW   modulator identification.
XX
OS   Chimeric - Homo sapiens.
XX
PN   WO20029033-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31074.
XX
PR   03-OCT-2000; 2000US-0679664.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI   Simin RT;
XX
DR   WPI; 2002-330170/36.
XX
PT   Novel G-protein fusion receptor, useful for identifying modulators of
PT   CaR, mGluR and GABABR, comprises G-protein joined to the intracellular
PT   domain of the receptor -
XX
PS   Example 3; Fig 12; 168pp; English.
XX
CC   The invention comprises G-protein fusion receptors - comprising
CC   extracellular, transmembrane and intracellular domains similar to CaR,
CC   mGluR or GABAB receptor sequences. The G-protein fusion receptors of the
CC   invention may also possess a linker joined to the carboxy terminus of the
CC   intracellular domain, and a G-protein joined to the linker. The G-protein
CC   fusion receptors of the invention are useful for identifying modulators
CC   of CaR, mGluR and GABABR for use in treating associated conditions. The
CC   present amino acid sequence was used in the production of the invention.
```

```
XX
SQ   Sequence 1394 AA;
XX
Query Match    98.2%;  Score 1827;  DB 23;  Length 1394;
Best Local Similarity  98.6%;  Pred. No. 4.9e-172;
Matches 353;  Conservative 2;  Mismatches 3;  Indels 0;  Gaps 0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKDRARRELKLLLGTGESGKSTFIKQMR 60
         ||||||||||||||||||||||||||||  ||||| |||||||||||||||||||||||
Db  1036 MTLESIMACCLSEEAKEARRINDEIERQLRDKDRDARRELKLLLGTGESGKSTFIKQMR 1095

Qy    61 IIHGSGCYSDEDKDRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1096 IIHGSGYSDEDKDRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 1155

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1156 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1215

Qy   181 RVRVPPTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1216 RVRVPPTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 1275

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1276 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1335

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1336 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL 1393

RESULT 9
AAY49134
ID   AAY49134 standard; Protein; 1397 AA.
XX
AC   AAY49134;
XX
DT   07-JAN-2000 (first entry)
XX
DE   pmGluR2/CaR*Galpha15+3Ala linker fusion construct protein sequence.
XX
KW   G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
KW   metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW   gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW   spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW   Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW   cognitive disorder.
XX
OS   Homo sapiens.
XX
PN   WO9951641-A1.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999; 99WO-US07333.
XX
PR   03-APR-1998; 98US-0080671.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Simin RT;
PI   Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
XX
DR   WPI; 1999-610995/52.
DR   N-PSDB; AAZ31065.
XX
PT   New G-protein fusion receptors and chimeras containing domains from
PT   different receptors, used to screen for modulators, potentially useful
PT   e.g. for treating or preventing stroke or Alzheimer's disease.
XX
PS   Example 1; Fig 12; 255pp; English.
XX
```

```
CC   The invention relates to G-protein fusion receptors (I) comprising:
CC   (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC   intracellular (ICD) domains, each chosen independently from a CaR
CC   (calcium receptor), GluR (metabotropic glutamate receptor) or GABABR
CC   (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC   the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC   linker. (I) and recombinant chimeric receptors (CR) without the GP
CC   component, are used to assess function of the various domains and to
CC   identify compounds (e.g. allosteric modulators or antagonists) that act
CC   on these domains. The modulators are potentially useful for treating or
CC   preventing diseases associated with the receptors, e.g. stroke, head or
CC   spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC   Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC   disorders and depression. Nucleic acid (II) that encodes (I) is used:
CC   (1) for recombinant production of corresponding proteins; and (2) to
CC   produce cells used in screening for modulators. Use of CaR and mGluR
CC   domains allows presentation of GABABR domains, to a binding agent, in a
CC   form more like the natural domain structure compared with use of
CC   incomplete receptors, lacking one or more domains. By shuffling different
CC   domains, agents can be identified that affect particular domains of a
CC   receptor.
```

```
XX
SQ   Sequence 1397 AA;
XX
Query Match    98.2%;  Score 1827;  DB 20;  Length 1397;
Best Local Similarity  98.6%;  Pred. No. 4.9e-172;
Matches 353;  Conservative 2;  Mismatches 3;  Indels 0;  Gaps 0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKDRARRELKLLLGTGESGKSTFIKQMR 60
         ||||||||||||||||||||||||||||  ||||| |||||||||||||||||||||||
Db  1039 MTLESIMACCLSEEAKEARRINDEIERQLRDKDRDARRELKLLLGTGESGKSTFIKQMR 1098

Qy    61 IIHGSGYSDEDKDRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1099 IIHGSGYSDEDKDRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 1158

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1159 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1218

Qy   181 RVRVPPTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1219 RVRVPPTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 1278

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1279 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1338

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db  1339 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL 1396

RESULT 10
AAO15100
ID   AAO15100 standard; Protein; 1397 AA.
XX
AC   AAO15100;
XX
DT   22-AUG-2002 (first entry)
XX
DE   Human pmGluR2-CaR*G-alpha-qi5+3Ala linker fusion construct protein.
XX
KW   G-protein fusion receptor; CaR; extracellular domain; mGluR; GABABR;
KW   transmembrane domain; intracellular domain; chimeric receptor; mGluR;
KW   modulator identification.
XX
OS   Human; G-protein fusion receptor; intracellular domain; CaR; mGluR;
OS   Chimeric - Homo sapiens.
OS   Chimeric - unidentified.
XX
PN   WO20029033-A2.
XX
PD   11-APR-2002.
```

XX
PF   03-OCT-2001; 2001WO-US31074.
XX
PR   03-OCT-2000; 2000US-0679664..
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI   Simin RT;
XX
DR   WPI; 2002-330170/36.
XX
PS   Disclosure; Fig 12; 168pp; English.
XX
PT   Novel G-protein fusion receptor, useful for identifying modulators of
PT   CaR, mGluR and GABABR, comprises G-protein joined to the intracellular
PT   domain of the receptor -
XX
CC   The invention comprises G-protein fusion receptors - comprising
CC   extracellular, transmembrane and intracellular domains similar to CaR,
CC   mGluR or GABAB receptor sequences. The G-protein fusion receptors of the
CC   invention may also possess a linker joined to the carboxy terminus of the
CC   intracellular domain, and a G-protein joined to the linker. The G-protein
CC   fusion receptors of the invention are useful for identifying modulators
CC   of CaR, mGluR and GABAB for use in treating associated conditions. The
CC   present amino acid sequence was used in the production of the invention.
XX
SQ   Sequence   1397 AA;

Query Match     98.2%;   Score 1827;   DB 23;   Length 1397;
Best Local Similarity   98.6%;   Pred. No. 4.9e-172;
Matches 353;   Conservative   2;   Mismatches   3;   Indels   0;   Gaps   0;

```
Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLILLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db   1039 MTLESIMACCLSEEAKEARRINDEIERQLRDKRDARRELKLILLGTGESGKSTFIKQMR 1098

Qy     61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 1158

Qy    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1218

Qy    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 1278

Qy    241 ESDNENRMEESKALFRTIITTYPWFQNSSVILFLNKKDLLEEKIMYSHLVYFPEYDGPQR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 ESDNENRMEESKALFRTIITTYPWFQNSSVILFLNKKDLLEEKIMYSHLVYFPEYDGPQR 1338

Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNL 358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||: |
Db   1339 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKDCGL 1396
```

RESULT 11
AAO15105
ID   AAO15105 standard; Protein; 1402 AA.
XX
AC   AAO15105;
XX
DT   22-AUG-2002 (first entry)
XX
DE   Human ph2SPMGluR3-CaR*AAA*Gg15 fusion construct protein sequence.
XX
KW   Human; G-protein fusion receptor; extracellular domain; CaR;
KW   transmembrane domain; intracellular domain; mGluR; GABABR;
KW   modulator identification.
XX
OS   Chimeric - Homo sapiens.

---

XX
PN   WO200229033-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31074.
XX
PR   03-OCT-2000; 2000US-0679664.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI   Simin RT;
XX
PA   (NPSP-) NPS PHARM INC.
XX
DR   WPI; 2002-330170/36.
DR   N-PSDB; AAL43286.
XX
PS   Disclosure; Fig 18; 168pp; English.
XX
PT   Novel G-protein fusion receptor, useful for identifying modulators of
PT   CaR, mGluR and GABABR, comprises G-protein joined to the intracellular
PT   domain of the receptor -
XX
CC   The invention comprises G-protein fusion receptors - comprising
CC   extracellular, transmembrane and intracellular domains similar to CaR,
CC   mGluR or GABAB receptor sequences. The G-protein fusion receptors of the
CC   invention may also possess a linker joined to the carboxy terminus of the
CC   intracellular domain, and a G-protein joined to the linker. The G-protein
CC   fusion receptors of the invention are useful for identifying modulators
CC   of CaR, mGluR and GABAB for use in treating associated conditions. The
CC   present amino acid sequence was used in the production of the invention.
XX
SQ   Sequence   1402 AA;

Query Match     98.2%;   Score 1827;   DB 23;   Length 1402;
Best Local Similarity   98.6%;   Pred. No. 4.9e-172;
Matches 353;   Conservative   2;   Mismatches   3;   Indels   0;   Gaps   0;

```
Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLILLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db   1044 MTLESIMACCLSEEAKEARRINDEIERQLRDKRDARRELKLILLGTGESGKSTFIKQMR 1103

Qy     61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1104 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 1163

Qy    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1164 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1223

Qy    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1224 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 1283

Qy    241 ESDNENRMEESKALFRTIITTYPWFQNSSVILFLNKKDLLEEKIMYSHLVYFPEYDGPQR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1284 ESDNENRMEESKALFRTIITTYPWFQNSSVILFLNKKDLLEEKIMYSHLVYFPEYDGPQR 1343

Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNL 358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||: |
Db   1344 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKDCGL 1401
```

RESULT 12
AAY49131
ID   AAY49131 standard; Protein; 1418 AA.
XX
AC   AAY49131;
XX
DT   07-JAN-2000 (first entry)
XX
DE   mGluR8/CaR*Galphaqi5 fusion construct protein sequence.
XX

KW   G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
KW   metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW   gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW   spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW   Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW   cognitive disorder.
XX
OS   Homo sapiens.
XX
PN   WO9951641-A1.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;  99WO-US07333.
XX
PR   03-APR-1998;  98US-0080671.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann TM, Hammerland LG,  Storjohann LL,  Busby JG,  Garrett JE;
PI   Simin RT;
XX
DR   WPI; 1999-610995/52.
DR   N-PSDB; AAZ31062.
XX
PS   Example 1; Fig 12; 255pp; English.
XX
PT   The invention relates to G-protein fusion receptors (I) comprising:
PT   (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
PT   intracellular (ICD) domains, each chosen independently from a CaR
PT   (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
PT   (gamma-aminobutyric acid receptor); (2) an optional linker attached to
PT   the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
PT   linker. (I), and recombinant chimeric receptors (CR) without the GP
PT   component, are used to assess function of the various domains and to
PT   identify compounds (e.g. allosteric modulators or antagonists) that act
PT   on these domains. The modulators are potentially useful for treating or
PT   preventing diseases associated with the receptors, e.g. stroke, head or
PT   spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
PT   disorders and depression. Nucleic acid (II) that encodes (I) is used:
PT   (1) for recombinant production of corresponding proteins; and (2) to
PT   produce cells used in screening for modulators. Use of CaR and mGluR
PT   domains allows presentation of GABABR domains, to a binding agent, in a
PT   form more like the natural domain structure compared with use of
PT   incomplete receptors, lacking one or more domains. By shuffling different
PT   domains, agents can be identified that affect particular domains of a
PT   receptor.
XX
SQ   Sequence   1418 AA;

Qy          Query Match             98.2%;  Score 1827;  DB 20;  Length 1418;
            Best Local Similarity   98.6%;
            Matches 353;  Conservative   2;  Mismatches   3;  Indels   0;  Gaps   0;

Qy        1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMR   60
             ||||||||||||||||||||||||:||||||||||:||||||||||||||||||||||
Db     1060  MTLESIMACCLSEEAKEARRINDEIERQLRDKRDARRELKLLLGTGESGKSTFIKQMR  1119

Qy       61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1120  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  1179

Qy      121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1180  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  1239

Qy      181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     1240  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  1299
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYPPEYDGPQR  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1300  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYPPEYDGPQR  1359

Qy      301  DAQQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNL  358
             ||||||||||||||||||||||||||||||||||||||||||||||||||:  ||
Db     1360  DAQQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKDCGL  1417

RESULT 13
ID   AAO15097
AC   AAO15097;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human mGluR8-CaR*G-alpha-qi5 fusion construct protein.
XX
OS   Chimeric - Homo sapiens.
XX
PN   WO20029033-A2.
XX
PD   11-APR-2002.
XX
DR   WPI; 2002-330170/36.
XX
PI   Stormann T, Hammerland LG,  Storjohann LL,  Busby JG,  Garrett JE;
PI   Simin RT;
XX
PA   (NPSP-) NPS PHARM INC.
XX
PR   03-OCT-2000;  2000US-0679664.
XX
PF   03-OCT-2001;  2001WO-US31074.
XX
PD   AAO15097 (standard);  Protein;  1418 AA.
ID   AAO15097
AAO15097
RESULT 13

Db     1240  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  1299
Qy      241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYPPEYDGPQR  300
Db     1300  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYPPEYDGPQR  1359

Qy      301  DAQQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNL  358
             ||||||||||||||||||||||||||||||||||||||||||||||||||:  |
Db     1360  DAQQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKDCGL  1417

Example 3; Fig 12; 168pp; English.
XX
PT   Novel G-protein fusion receptor, useful for identifying modulators of
PT   CaR, mGluR and GABABR, comprises G-protein joined to the intracellular
PT   domain of the receptor -
XX
CC   The invention comprises G-protein fusion receptors - comprising
CC   extracellular, transmembrane and intracellular domains similar to CaR,
CC   mGluR or GABABR receptor sequences. The G-protein fusion receptors of the
CC   invention may also possess a linker joined to the carboxy terminus of the
CC   intracellular domain, and a G-protein joined to the linker. The G-protein
CC   fusion receptors of the invention are useful for identifying modulators
CC   of CaR, mGluR and GABABR for use in treating associated conditions. The
CC   present amino acid sequence was used in the production of the invention.
XX
SQ   Sequence   1418 AA;

Qy          Query Match             98.2%;  Score 1827;  DB 23;  Length 1418;
            Best Local Similarity   98.6%;
            Matches 353;  Conservative   2;  Mismatches   3;  Indels   0;  Gaps   0;

Qy        1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMR   60
             ||||||||||||||||||||||||:||||||||||:||||||||||||||||||||||
Db     1060  MTLESIMACCLSEEAKEARRINDEIEROLRDKRDARRELKLLLGTGESGKSTFIKQMR  1119

Qy       61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1120  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  1179

Qy      121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

## RESULT 14

```
AA015103
ID   AA015103 standard; protein; 1421 AA.
XX
AC   AA015103;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human phmGluR8-CaR*AAA*G-alpha-qi5 fusion protein.
XX
KW   Human; G-protein fusion receptor; extracellular domain;
KW   transmembrane domain; intracellular domain; CaR; mGluR; GABABR;
KW   modulator identification.
XX
OS   Chimeric - Homo sapiens.
XX
XX
PN   WO20029033-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31074.
PR   03-OCT-2000; 2000US-0679664.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann T,  Hammerland LG,  Storjohann LL,  Busby JG,  Garrett JE;
PI   Simin RT;
XX
DR   WPI; 2002-330170/36.
XX
PT   Novel G-protein fusion receptor, useful for identifying modulators of
PT   CaR, mGluR and GABABR, comprises G-protein joined to the intracellular
PT   domain of the receptor -
XX
PS   Disclosure; Fig 16; 168pp; English.
XX
CC   The invention comprises G-protein fusion receptors - comprising
CC   extracellular, transmembrane and intracellular domains similar to CaR,
CC   mGluR or GABAB receptor sequences. The G-protein fusion receptors of the
CC   invention may also possess a linker joined to the carboxy terminus of the
CC   intracellular domain, and a G-protein joined to the linker. The G-protein
CC   fusion receptors of the invention are useful for identifying modulators
CC   of CaR, mGluR and GABABR for use in treating associated conditions. The
CC   present amino acid sequence was used in the production of the invention.
XX
SQ   Sequence  1421 AA;

Query Match             98.2%;  Score 1827;  DB 23;  Length 1421;
Best Local Similarity   98.6%;
Matches  353;  Conservative   2;  Mismatches   3;  Indels    0;  Gaps    0;

Qy           1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
               |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        1063 MTLESIMACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 1122

Qy          61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKEHNKAHAQLVREVDVEK 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1123 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKEHNKAHAQLVREVDVEK 1182

Qy         181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLV 240
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1183 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1242

Qy         241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1303 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1362

Qy         301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
               |||||||||:||||||||||||||||||||||||||||||||||||||||||||| |
Db        1363 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL 1420
```

## RESULT 15

```
AA015102
ID   AA015102 standard; protein; 1422 AA.
XX
AC   AA015102;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human phmGluR4-CaR*AAA*G-alpha-qi5 fusion protein.
XX
KW   Human; G-protein fusion receptor; extracellular domain;
KW   transmembrane domain; intracellular domain; CaR; mGluR; GABABR;
KW   modulator identification.
XX
OS   Chimeric - Homo sapiens.
XX
XX
PN   WO20029033-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31074.
PR   03-OCT-2000; 2000US-0679664.
XX
PA   (NPSP-) NPS PHARM INC.
XX
PI   Stormann T,  Hammerland LG,  Storjohann LL,  Busby JG,  Garrett JE;
PI   Simin RT;
XX
DR   WPI; 2002-330170/36.
XX
PT   Novel G-protein fusion receptor, useful for identifying modulators of
PT   CaR, mGluR and GABABR, comprises G-protein joined to the intracellular
PT   domain of the receptor -
XX
PS   Disclosure; Fig 16; 168pp; English.
XX
CC   The invention comprises G-protein fusion receptors - comprising
CC   extracellular, transmembrane and intracellular domains similar to CaR,
CC   mGluR or GABAB receptor sequences. The G-protein fusion receptors of the
CC   invention may also possess a linker joined to the carboxy terminus of the
CC   intracellular domain, and a G-protein joined to the linker. The G-protein
CC   fusion receptors of the invention are useful for identifying modulators
CC   of CaR, mGluR and GABABR for use in treating associated conditions. The
CC   present amino acid sequence was used in the production of the invention.
XX
SQ   Sequence  1422 AA;

Query Match             98.2%;  Score 1827;  DB 23;  Length 1422;
Best Local Similarity   98.6%;
Matches  353;  Conservative   2;  Mismatches   3;  Indels    0;  Gaps    0;

Qy           1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
               |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        1123 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKEHNKAHAQLVREVDVEK 1182

Qy         121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1183 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 1242

Qy         181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLV 240
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1243 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLV 1302

Qy         241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1303 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 1362

Qy         301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
               |||||||||:||||||||||||||||||||||||||||||||||||||||||||| |
Db        1363 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL 1420
```

```
Db   1064  MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGIGESGKSTFIKQMR  1123

Qy     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  1183

Qy    121  VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184  VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  1243

Qy    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  1303

Qy    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  1363

Qy    301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL    358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:    |
Db   1364  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKDCGL   1421
```

Search completed: June 30, 2003, 16:16:16
Job time : 45 secs

OM protein - protein search, using sw model

```
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:         June 30, 2003, 16:15:31 ; Search time 15 seconds
                                           (without alignments)
                                           704.189 Million cell updates/sec

Title:          US-09-899-295-2
Perfect score:  1860
Sequence:       1 MTLESIMACCLSEEAKEARR..............VFAAVKDTILQLNKEYNLV 359
Scoring table:  BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:       262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:    262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Issued_Patents_AA:*
        1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
        2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
        3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
        4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
        5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
        6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1066 | 57.3 | 374 | 3 | US-08-878-801-4 | Sequence 4, Appli |
| 2 | 1063 | 57.2 | 374 | 4 | US-09-442-349A-29 | Sequence 29, Appl |
| 3 | 1063 | 57.2 | 374 | 4 | US-09-442-349A-31 | Sequence 31, Appl |
| 4 | 1062 | 57.1 | 374 | 4 | US-09-442-349A-21 | Sequence 21, Appl |
| 5 | 1062 | 57.1 | 374 | 4 | US-09-442-349A-23 | Sequence 23, Appl |
| 6 | 1062 | 57.1 | 374 | 4 | US-09-442-349A-25 | Sequence 25, Appl |
| 7 | 1062 | 57.1 | 374 | 4 | US-09-442-349A-30 | Sequence 30, Appl |
| 8 | 1062 | 57.1 | 374 | 4 | US-09-442-349A-32 | Sequence 32, Appl |
| 9 | 1062 | 57.1 | 374 | 4 | US-09-442-349A-34 | Sequence 34, Appl |
| 10 | 1061 | 57.0 | 374 | 4 | US-09-442-349A-2 | Sequence 2, Appli |
| 11 | 1061 | 57.0 | 374 | 4 | US-09-442-349A-22 | Sequence 22, Appl |
| 12 | 1061 | 57.0 | 374 | 4 | US-09-442-349A-24 | Sequence 24, Appl |
| 13 | 1061 | 57.0 | 374 | 4 | US-09-442-349A-33 | Sequence 33, Appl |
| 14 | 1058 | 56.9 | 374 | 4 | US-09-442-349A-26 | Sequence 26, Appl |
| 15 | 1058 | 56.9 | 374 | 4 | US-09-442-349A-28 | Sequence 28, Appl |
| 16 | 1057 | 56.8 | 374 | 4 | US-09-442-349A-20 | Sequence 20, Appl |
| 17 | 1057 | 56.8 | 374 | 4 | US-09-442-349A-27 | Sequence 27, Appl |
| 18 | 1055 | 56.7 | 374 | 4 | US-09-218-489-2 | Sequence 2, Appll |
| 19 | 1055 | 56.7 | 374 | 4 | US-09-442-349A-14 | Sequence 14, Appl |
| 20 | 1054 | 56.7 | 374 | 4 | US-09-442-349A-16 | Sequence 16, Appl |
| 21 | 1053 | 56.6 | 365 | 4 | US-09-442-349A-108 | Sequence 108, App |
| 22 | 1053 | 56.6 | 374 | 4 | US-09-442-349A-6 | Sequence 6, Appli |
| 23 | 1053 | 56.6 | 374 | 4 | US-09-442-349A-8 | Sequence 8, Appli |
| 24 | 1053 | 56.6 | 374 | 4 | US-09-442-349A-10 | Sequence 10, Appl |
| 25 | 1053 | 56.6 | 374 | 4 | US-09-442-349A-15 | Sequence 15, Appl |
| 26 | 1053 | 56.6 | 374 | 4 | US-09-442-349A-17 | Sequence 17, Appl |
| 27 | 1053 | 56.6 | 374 | 4 | US-09-442-349A-19 | Sequence 19, Appl |
| 28 | 1052 | 56.6 | 374 | 4 | US-09-442-349A-1 | Sequence 1, Appli |
| 29 | 1052 | 56.6 | 374 | 4 | US-09-442-349A-7 | Sequence 7, Appli |
| 30 | 1052 | 56.6 | 374 | 4 | US-09-442-349A-9 | Sequence 9, Appli |
| 31 | 1052 | 56.6 | 374 | 4 | US-09-442-349A-18 | Sequence 18, Appl |
| 32 | 1049 | 56.4 | 374 | 4 | US-09-442-349A-11 | Sequence 11, Appl |
| 33 | 1049 | 56.4 | 374 | 4 | US-09-442-349A-13 | Sequence 13, Appl |
| 34 | 1048 | 56.3 | 374 | 4 | US-09-442-349A-5 | Sequence 5, Appll |
| 35 | 1048 | 56.3 | 374 | 4 | US-09-442-349A-12 | Sequence 12, Appl |
| 36 | 1046 | 56.2 | 374 | 4 | US-09-442-349A-36 | Sequence 36, Appl |
| 37 | 1046 | 56.2 | 374 | 4 | US-09-442-349A-38 | Sequence 38, Appl |
| 38 | 1045 | 56.2 | 374 | 4 | US-09-442-349A-3 | Sequence 3, Appli |
| 39 | 1045 | 56.2 | 374 | 4 | US-09-442-349A-37 | Sequence 37, Appl |
| 40 | 1045 | 56.2 | 374 | 4 | US-09-442-349A-48 | Sequence 48, Appl |
| 41 | 1045 | 56.2 | 374 | 4 | US-09-442-349A-50 | Sequence 50, Appl |
| 42 | 1045 | 56.2 | 374 | 4 | US-09-442-349A-54 | Sequence 54, Appl |
| 43 | 1045 | 56.2 | 374 | 4 | US-09-442-349A-56 | Sequence 56, Appl |
| 44 | 1044 | 56.1 | 374 | 3 | US-08-878-801-2 | Sequence 2, Appl |
| 45 | 1044 | 56.1 | 374 | 4 | US-09-442-349A-39 | Sequence 39, Appl |

ALIGNMENTS

```
RESULT 1
US-08-878-801-4
  Sequence 4, Application US/08878801
  Patent No. 6004808
GENERAL INFORMATION:
  APPLICANT: Negulescu, Paul
  APPLICANT: Offermanns, Stefan
  APPLICANT: Simon, Melvin
  APPLICANT: Zuker, Charles
  TITLE OF INVENTION: PROMISCUOUS G-PROTEINS COMPOSITIONS AND THEIR USE
  FILE REFERENCE: 08366/002001
  CURRENT APPLICATION NUMBER: US-09/878,801
  CURRENT FILING DATE: 1997-06-19
  EARLIER APPLICATION NUMBER: US 60/020,234
  EARLIER FILING DATE: 1996-06-21
  NUMBER OF SEQ ID NOS: 4
  SOFTWARE: FastSEQ for Windows Version 3.0
  SEQ ID NO 4
  LENGTH: 374
  TYPE: PRT
  ORGANISM: Homo sapiens
US-08-878-801-4

  Query Match             57.3%;  Score 1066;  DB 3;  Length 374;
  Best Local Similarity   57.3%;  Pred. No. 7.7e-100;
  Matches  207; Conservative  55; Mismatches  88; Indels  12; Gaps  2;
```

```
Qy     10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLILILGTGESGKSTFIKQMRIIHGSGYSD 69
          ||||||||||||| || ||||||||||||||||||||||||||||||||||||||||||||
Db     13 CLTEEEKTAARIDQEINRILLEQKKOEREELKLILLGPGESGKSTFIKQMRIIHGVGYSE 72

Qy     70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
          || ||| | ||||| ||||||| |||| ||:||| || ||| | |||| ||| || |||
Db     73 EDRRAFRLLIYQNIFVSMQAMIDAMDRLQIPFSRPDSKQHASLVMTQDPYKVSTFEKPYA 132

Qy    130 DAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
          |:| |||| || ||:|||||||: ||:|||||:||| | |||| ||:: |||:||||| ||
Db    133 VAMQYLWRDAGIRACYERRREFHLLDSAVYYLSHLERISEDSYIPTAQDVLRSRMPTTGI 192

Qy    190 IEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
          ||||| | |||||||||||||||| ||||||||||||| ||||||||||||||||||||
Db    193 NEYCEFSVKKTKLRIVDVGGQRSERRKWIHCFENVTALIYLASLSEYDQCLEENDQENRME 252

Qy    250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
          |||||||:|:|||||:|||::||||||||||||:||:||||:|::|::||||:::::|::
Db    253 ESLALFSTILELPWFKSTSVILFLNKTDILEDKIHTSHLATYFPSFQGPRRDAEAAKSFI 312

Qy    310 LKMFEVDL------NPDSD------KIIYSHFTCATDTENRFVFAAVKDTILQLNLKEYN 357
          ||:|||| | | |::|||||||||||||:: |||:||||||| |
Db    313 LDMYARVYASCAEPQDGGRKGSRARFFAHFTCATDTQSVRSVFKDVRDSVLARYLDEIN 372
```

```
Qy    358 LV 359
          |:
Db    373 LL 374
```

RESULT 2
US-09-442-349A-29

; Sequence 29, Application US/0942349A
; Patent No. 6462178
; GENERAL INFORMATION:
;   APPLICANT: Wong, Yung H
;   TITLE OF INVENTION: G Protein
;   FILE REFERENCE: M99/0101/US
;   CURRENT APPLICATION NUMBER: US/09/442,349A
;   CURRENT FILING DATE: 1999-11-17
;   NUMBER OF SEQ ID NOS: 116
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 29
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-29

```
Query Match      57.2%;  Score 1063;  DB 4;  Length 374;
Best Local Similarity   57.1%;  Pred. No. 1.6e-99;
Matches 206;  Conservative  57;  Mismatches 86;  Indels  12;  Gaps  2;

Qy   10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
        |:||:| | |:: :|| |:: |||||:|||:||||| |||||||||||||||||||||:||:
Db   13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy   70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
        ::|||  |||  ||| || :| || ||||| |||||:|| ||| ||:| ||::| |||||
Db   73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy  130 DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
        |:|:|:| | ::| |:  : :| :||:|||||:|:|:|||||| ||:| |||:|:|||||
Db  133 AAMQWLWRDAGIRACYERRREFHLLDSAVVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy  190 IEYPFDLQSVIFRMVDVGGQRSERRKKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
        ||| ||||||| ||||||||:|||||||||||||||:|||||||||||||||| ||| |||
Db  193 NEYCFSVQKTNLRIVDVGGQKSERRKWIHCFENVIALIYLASLSEYDQCLEENQENRMK 252

Qy  250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
        | |||||||||||||:|:|:||||||||||||:||| ||:|||:|||||:|||  :|||
Db  253 ESLALFGTIIELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPKQDAEAAKRFI 312

Qy  310 LKMFVDL-------NPDSD----KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
        |:|:                   |:|||||:||||||:||||||:||:|| ||||||:|
Db  313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCAIDTKNIQFVFEAVTDVIIQNNLKYIG 372

Qy  358 357
Db  373 L 373
```

RESULT 3
US-09-442-349A-31

; Sequence 31, Application US/0942349A
; Patent No. 6462178
; GENERAL INFORMATION:
;   APPLICANT: Wong, Yung H
;   TITLE OF INVENTION: G Protein
;   FILE REFERENCE: M99/0101/US
;   CURRENT APPLICATION NUMBER: US/09/442,349A
;   CURRENT FILING DATE: 1999-11-17
;   NUMBER OF SEQ ID NOS: 116
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 31

---

LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-31

```
Query Match      57.2%;  Score 1063;  DB 4;  Length 374;
Best Local Similarity   57.1%;  Pred. No. 1.6e-99;
Matches 206;  Conservative  57;  Mismatches 86;  Indels  12;  Gaps  2;

Qy   10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
        |:||:| | |:: :|| |:: |||||:|||:||||| |||||||||||||||||||||:||:
Db   13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy   70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
        ::|||  |||  ||| || :| || ||||| |||||:|| ||| ||:| ||::| |||||
Db   73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132
```

RESULT 4
US-09-442-349A-21

; Sequence 21, Application US/0942349A
; Patent No. 6462178
; GENERAL INFORMATION:
;   APPLICANT: Wong, Yung H
;   TITLE OF INVENTION: G Protein
;   FILE REFERENCE: M99/0101/US
;   CURRENT APPLICATION NUMBER: US/09/442,349A
;   CURRENT FILING DATE: 1999-11-17
;   NUMBER OF SEQ ID NOS: 116
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 374
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: chimera
US-09-442-349A-21

```
Query Match      57.2%;  Score 1062;  DB 4;  Length 374;
Best Local Similarity   57.1%;  Pred. No. 2e-99;
Matches 206;  Conservative  57;  Mismatches 86;  Indels  12;  Gaps  2;

Qy   10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
        |:||:| | |:: :|| |:: |||||:|||:||||| |||||||||||||||||||||:||:
Db   13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy   70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
        ::|||  |||  ||| || :| || ||||| |||||:|| ||| ||:| ||::| |||||
Db   73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy  130 DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
        |:|:|:| | ::| |:  : :| :||:|||||:|:|:|||||| ||:| |||:|:|||||
Db  133 AAMQWLWRDAGIRACYERRREFHLLDSAVVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy  190 IEYPFDLQSVIFRMVDVGGQRSERRKKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
        ||| ||||||| ||||||||:|||||||||||||||:|||||||||||||||| ||| |||
Db  193 NEYCFSVQKTNLRIVDVGGQKSERRKWIHCFENVIALIYLASLSEYDQCLEENQENRMK 252

Qy  250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
        | |||||||||||||:|:|:||||||||||||:||| ||:|||:|||||:|||  :|||
Db  253 ESLALFGTIIELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPKQDAEAAKRFI 312

Qy  310 LKMFVDL-------NPDSD----KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
        |:|:                   |:|||||:||||||:||||||:||:|| ||||||:|
Db  313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCAIDTKNIQFVFEAVTDVIIQNNLKYIG 372

Qy  358 L 358
        |
Db  373 L 373

Qy   70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
        ::|||  |||  ||| || :| || ||||| |||||:|| ||| ||:| ||::| |||||
Db   73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132
```

RESULT 5
US-09-442-349A-23
Sequence 23, Application US/0944234 9A
Patent No. 6462178
GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 23
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-23

Query Match      57.1%;   Score 1062;   DB 4;   Length 374;
Best Local Similarity   57.1%;   Pred. No. 2e-99;
Matches 206;  Conservative 56;  Mismatches 87;  Indels 12;  Gaps 2;

```
Qy   10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
          || :|: ||| ::|: ||| | ||::|::|||||||| ||||||||||||||||| ||::
Db   13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy   70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPV 129
          |::|:| ||||||||::::|||||::|:|:| | |::|::|| |:|:|:|| |:||:
Db   73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy  130 DAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
          :|: | |||:||| |:| ||||| :||: ||||:|:|| ::||  ||||||: ||||
Db  133 AAMQWLWRDAGIRACYERRREFHLLDSAVYYLSHLERITEEGVPTAQDVLRSRMPTTGI 192

Qy  190 IEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
          ||::|| ||||||| ||||||:||||:||||||||||| |||||||||| || :| |:|
Db  193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy  250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
          ||:|| ||| :|||:|:||||||||:| ||||| |||| ||||| :||:|||:||||:||
Db  253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPKQDAEAAKRFI 312

Qy  310 LKMFVDL------NPDSD------KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
          | |: :|      ||::       | |||||||||||||::| |||| |:||:|||:||||
Db  313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTSNIQFVFEAVTDVIIQNNLKYIG 372

Qy  358 L 358
          |
Db  373 L 373
```

RESULT 6
US-09-442-349A-25
Sequence 25, Application US/0944234 9A
Patent No. 6462178
GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 25
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
; OTHER INFORMATION: Description of Artificial Sequence: chimera
US-09-442-349A-25

Query Match      57.1%;   Score 1062;   DB 4;   Length 374;
Best Local Similarity   57.1%;   Pred. No. 2e-99;
Matches 206;  Conservative 56;  Mismatches 87;  Indels 12;  Gaps 2;

```
Qy   10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
          || :|: ||| ::|: ||| | ||::|::|||||||| ||||||||||||||||| ||::
Db   13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy   70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPV 129
          |::|:| ||||||||::::|||||::|:|:| | |::|::|| |:|:|:|| |:||:
Db   73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy  130 DAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
          :|: | |||:||| |:| ||||| :||: ||||:|:|| ::||  ||||||: ||||
Db  133 AAMQWLWRDAGIRACYERRREFHLLDSAVYYLSHLERITEEGVPTAQDVLRSRMPTTGI 192

Qy  190 IEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
          ||::|| ||||||| ||||||:||||:||||||||||| |||||||||| || :| |:|
Db  193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy  250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
          ||:|| ||| :|||:|:||||||||:| ||||| |||| ||||| :||:|||:||||:||
Db  253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPKQDAEAAKRFI 312

Qy  310 LKMFVDL------NPDSD------KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
          | |: :|      ||::       | |||||||||||||::| |||| |:||:|||:||||
Db  313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTSNIQFVFEAVTDVIIQNNLKYIG 372

Qy  358 L 358
          |
Db  373 L 373
```

RESULT 7
US-09-442-349A-30
Sequence 30, Application US/0944234 9A
Patent No. 6462178
GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 30
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence

```
;  FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: G protein
;    OTHER INFORMATION: chimera
US-09-442-349A-30

Query Match   57.1%;  Score 1062;  DB 4;  Length 374;
Best Local Similarity   57.1%;  Pred. No. 2e-99;
Matches 206;  Conservative  57;  Mismatches  86;  Indels  12;  Gaps   2;

Qy    10 CLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
            |:|:| |: |:| | |   |||||||||||||||||||||||||||||||||||||||
Db    13 CLTEDEKAAARVDQEINRILLEOKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy    70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPVV 129
            |:||| | ||| |||||| |||||| ||||||   |||:|||||  |:  ||| ::|
Db    73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy   130 DAIKSLWNDPGIQECYDRREYQLSDSTKYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
           |:: ||| |||||||||||||||||||||||||||||||||||: ||| | |||||||
Db   133 AAMQWLWRDAGIRACYERRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy   190 IEYPFDLQSVIFRMVDVGGQRSERRKKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
            |:| ||||||||||||||||||||||||||||||||||||||::||||||||||::
Db   193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy   250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
            ||||| || | ||||||||||||||||||||||||||||||||||||||| |||||||
Db   253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPRKQDAEAAKRFI 312

Qy   310 LKMFVDL------NPDSD-----KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
           ||| ||:          ||| |:  :||||| |||||||||||||| ||| |||||||
Db   313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTKNIQFVFEAVTDVIIQNNLKYIG 372

Qy   358 L 358
           |
Db   373 L 373

RESULT 8
US-09-442-349A-32
;  Sequence 32, Application US/0944234 9A
;  Patent No. 6462178
; GENERAL INFORMATION:
;  APPLICANT:  Wong, Yung H
;  TITLE OF INVENTION:  G Protein
;  FILE REFERENCE:  M99/0101/US
;  CURRENT APPLICATION NUMBER:  US/09/442,349A
;  CURRENT FILING DATE:  1999-11-17
;  NUMBER OF SEQ ID NOS:  116
;  SOFTWARE:  PatentIn Ver. 2.1
;  SEQ ID NO 32
;  LENGTH: 374
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-32

Query Match   57.1%;  Score 1062;  DB 4;  Length 374;
Best Local Similarity   57.1%;  Pred. No. 2e-99;
Matches 206;  Conservative  56;  Mismatches  87;  Indels  12;  Gaps   2;

Qy    10 CLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
            |:|:| |: |:| | |   |||||||||||||||||||||||||||||||||||||||
Db    13 CLTEDEKAAARVDQEINRILLEOKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy    70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPVV 129
            |:||| | ||| |||||| |||||| ||||||   |||:|||||  |:  ||| ::|
Db    73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy   130 DAIKSLWNDPGIQECYDRREYQLSDSTKYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
           |:: ||| |||||||||||||||||||||||||||||||||||: ||| | |||||||
Db   133 AAMQWLWRDAGIRACYERRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy   190 IEYPFDLQSVIFRMVDVGGQRSERRKKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
            |:| ||||||||||||||||||||||||||||||||||||||::||||||||||::
Db   193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy   250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
            ||||| || | ||||||||||||||||||||||||||||||||||||||| |||||||
Db   253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPRKQDAEAAKRFI 312

Qy   310 LKMFVDL------NPDSD-----KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
           ||| ||:          ||| |:  :||||| |||||||||||||| ||| |||||||
Db   313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTKNIQFVFEAVTDVIIQNNLKYIG 372

Qy   358 L 358
           |
Db   373 L 373
```

```
Db   133 AAMQWLWRDAGIRACYERRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy   190 IEYPFDLQSVIFRMVDVGGQRSERRKKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
            |:| ||||||||||||||||||||||||||||||||||||||::||||||||||::
Db   193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy   250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
            ||||| || | ||||||||||||||||||||||||||||||||||||||| |||||||
Db   253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPRKQDAEAAKRFI 312

Qy   310 LKMFVDL------NPDSD-----KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
           ||| ||:          ||| |:  :||||| |||||||||||||| ||| |||||||
Db   313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTKNIQFVFEAVTDVIIQNNLKYIG 372

Qy   358 L 358
           |
Db   373 L 373

RESULT 9
US-09-442-349A-34
;  Sequence 34, Application US/0944234 9A
;  Patent No. 6462178
; GENERAL INFORMATION:
;  APPLICANT:  Wong, Yung H
;  TITLE OF INVENTION:  G Protein
;  FILE REFERENCE:  M99/0101/US
;  CURRENT APPLICATION NUMBER:  US/09/442,349A
;  CURRENT FILING DATE:  1999-11-17
;  NUMBER OF SEQ ID NOS:  116
;  SOFTWARE:  PatentIn Ver. 2.1
;  SEQ ID NO 34
;  LENGTH: 374
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: chimera
US-09-442-349A-34

Query Match   57.1%;  Score 1062;  DB 4;  Length 374;
Best Local Similarity   57.1%;  Pred. No. 2e-99;
Matches 206;  Conservative  56;  Mismatches  87;  Indels  12;  Gaps   2;

Qy    10 CLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
            |:|:| |: |:| | |   |||||||||||||||||||||||||||||||||||||||
Db    13 CLTEDEKAAARVDQEINRILLEOKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy    70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPVV 129
            |:||| | ||| |||||| |||||| ||||||   |||:|||||  |:  ||| ::|
Db    73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy   130 DAIKSLWNDPGIQECYDRREYQLSDSTKYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
           |:: ||| |||||||||||||||||||||||||||||||||||: ||| | |||||||
Db   133 AAMQWLWRDAGIRACYERRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy   190 IEYPFDLQSVIFRMVDVGGQRSERRKKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
            |:| ||||||||||||||||||||||||||||||||||||||::||||||||||::
Db   193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy   250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
            ||||| || | ||||||||||||||||||||||||||||||||||||||| |||||||
Db   253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPRKQDAEAAKRFI 312

Qy   310 LKMFVDL------NPDSD-----KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
           ||| ||:          ||| |:  :||||| |||||||||||||| ||| |||||||
Db   313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTKNIQFVFEAVTDVIIQNNLKYIG 372

Qy   358 L 358
           |
Db   373 L 373
```

RESULT 10
US-09-442-349A-2

; Sequence 2, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-2

  Query Match             57.0%;  Score 1061; DB 4;    Length 374;
  Best Local Similarity   57.1%;  Pred. No. 2.5e-99;
  Matches  206; Conservative  56; Mismatches  87;  Indels  12;  Gaps  2;

Qy      10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
           |::|:|:|:|||:::||::::||||::::|::|:|||||||||||||||||:::|:::
Db      13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy      70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
           |::::|| |||||||||:|:|||:||||||| ||:::|| |:|| ||||:||||:|||||
Db      73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy     130 DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
           :|::||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db     133 AAMQWLWRDAGIRACYERRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy     190 IEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
           ::|||||:|||||:||||||||||||||||||||:||:|:||||:|||||||||:|:|||
Db     193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy     250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
           :|:|:||||||:|||||||||||||||||:|||||:||||||||||:|::|| ||||||
Db     253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPRQDAEAAKRFI 312

Qy     310 LKMFVDL-----NPDSD--------KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
           ||:|||:       |:|        |||||||||||||:|||::|||||||:||::||:||
Db     313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTNNIQFVFDAVTDVIIQNNLKYIG 372

Qy     358 L 358
           |
Db     373 L 373

─────────────────────────────────────────────

RESULT 11
US-09-442-349A-22

; Sequence 22, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 22
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-22

  Query Match             57.0%;  Score 1061; DB 4;    Length 374;
  Best Local Similarity   57.1%;  Pred. No. 2.5e-99;
  Matches  206; Conservative  56; Mismatches  87;  Indels  12;  Gaps  2;

Qy      10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
           |::|:|:|:|||:::||::::||||::::|::|:|||||||||||||||||:::|:::
Db      13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy      70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
           |::::|| |||||||||:|:|||:||||||| ||:::|| |:|| ||||:||||:|||||
Db      73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy     130 DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
           :|::||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db     133 AAMQWLWRDAGIRACYERRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy     190 IEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
           ::|||||:|||||:||||||||||||||||||||:||:|:||||:|||||||||:|:|||
Db     193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy     250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
           :|:|:||||||:|||||||||||||||||:|||||:||||||||||:|::|| ||||||
Db     253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPRQDAEAAKRFI 312

Qy     310 LKMFVDL-----NPDSD--------KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
           ||:|||:       |:|        |||||||||||||:|||::|||||||:||::||:||
Db     313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTNNIQFVFDAVTDVIIQNNLKYIG 372

Qy     358 L 358
           |
Db     373 L 373

─────────────────────────────────────────────

RESULT 12
US-09-442-349A-24

; Sequence 24, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Wong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 24
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: chimera
US-09-442-349A-24

  Query Match             57.0%;  Score 1061; DB 4;    Length 374;
  Best Local Similarity   57.1%;  Pred. No. 2.5e-99;
  Matches  206; Conservative  56; Mismatches  87;  Indels  12;  Gaps  2;

Qy      10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
           |::|:|:|:|||:::||::::||||::::|::|:|||||||||||||||||:::|:::
Db      13 CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE 72

Qy      70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
           |::::|| |||||||||:|:|||:||||||| ||:::|| |:|| ||||:||||:|||||
Db      73 EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA 132

Qy     130 DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
           :|::||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db     133 AAMQWLWRDAGIRACYERRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPTTGI 192

Qy     190 IEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
           ::|||||:|||||:||||||||||||||||||||:||:|:||||:|||||||||:|:|||
Db     193 NEYCFSVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK 252

Qy     250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
           :|:|:||||||:|||||||||||||||||:|||||:||||||||||:|::|| ||||||
Db     253 ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPRQDAEAAKRFI 312

Qy     310 LKMFVDL-----NPDSD--------KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
           ||:|||:       |:|        |||||||||||||:|||::|||||||:||::||:||
Db     313 LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTNNIQFVFDAVTDVIIQNNLKYIG 372

Qy     358 L 358
           |
Db     373 L 373

```
RESULT 13
US-09-442-349A-33
Sequence 33, Application US/0944234 9A
Patent No. 6462178
GENERAL INFORMATION:
   APPLICANT: Wong, Yung H
   TITLE OF INVENTION: G Protein
   FILE REFERENCE: M99/0101/US
   CURRENT APPLICATION NUMBER: US/09/442,349A
   CURRENT FILING DATE: 1999-11-17
   NUMBER OF SEQ ID NOS: 116
   SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 33
   LENGTH: 374
   TYPE: PRT
   ORGANISM: Artificial Sequence
   FEATURE:
   OTHER INFORMATION: Description of Artificial Sequence: G protein
   OTHER INFORMATION: chimera
US-09-442-349A-33

  Query Match             57.0%;  Score 1061;  DB 4;  Length 374;
  Best Local Similarity   57.1%;  Pred. No. 2.5e-99;
  Matches   206;  Conservative   56;  Mismatches   87;  Indels   12;  Gaps    2;

Qy    10  CLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD   69
          ||:::||| ||:  |::|  |:: ||:|||:|||| |:|||||||||||||||||||| ||:
Db    13  CLTEDEKAAARVDQEINRILLEQKKQDRGELKLILLGPGESGKSTFIKQMRIIHGAGYSE   72

Qy    70  EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV  129
          |::|||  |||||||.:|:|||:|||::|:|:|: |::.::|:|| ::| |||||||||||
Db    73  EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA  132

Qy   130  DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI  189
          :|:. || .| || |||||: ||||||||||||||||||||||||||||||||||||||||
Db   133  AAMQWLWRDAGIRACYERRREFHLLDSAVYYLSHLERITEEGYVPTAQDVLRSRMPTTGI  192

Qy   190  IEYPFDLQSVIFRMVDVGGQQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME  249
          :|| || ::: ||||||||||:||||||||||||||:::: :|||||||.|| |::|||||
Db   193  NEYCFSVQKTNLRIVDVGGQQKSERRKWIHCFENVIALILASLSEYDQCLEENNQENRMK  252

Qy   250  ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI  309
          |||||| || | |||:::||||||||.|:|||||| |||| |||::|||:||| ||:||
Db   253  ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPKQDAEAAKRFI  312

Qy   310  LKMFVDL------NPDSD------KIIYSHFTCATDTENIREVFAAVKDTILQLNLKEYN  357
          | |:   |||||   |:: ||||||:||||||||||||| ||| ||:| |::|:||| :|
Db   313  LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTCNIQFVFNAVTDVIIQNNLKYIG  372

Qy   358  L  358
          |
Db   373  L  373
```

```
RESULT 14
US-09-442-349A-26
; Sequence 26, Application US/0944349A
```

```
RESULT 15
US-09-442-349A-28
Sequence 28, Application US/0944234 9A
Patent No. 6462178
GENERAL INFORMATION:
   APPLICANT: Wong, Yung H
   TITLE OF INVENTION: G Protein
   FILE REFERENCE: M99/0101/US
   CURRENT APPLICATION NUMBER: US/09/442,349A
   CURRENT FILING DATE: 1999-11-17
   NUMBER OF SEQ ID NOS: 116
   SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
   LENGTH: 374
   TYPE: PRT
   ORGANISM: Artificial Sequence
   FEATURE:
   OTHER INFORMATION: Description of Artificial Sequence: G protein
   OTHER INFORMATION: chimera
US-09-442-349A-28

  Query Match             56.9%;  Score 1058;  DB 4;  Length 374;
```

```
Patent No. 6462178
GENERAL INFORMATION:
   APPLICANT: Wong, Yung H
   TITLE OF INVENTION: G Protein
   FILE REFERENCE: M99/0101/US
   CURRENT APPLICATION NUMBER: US/09/442,349A
   CURRENT FILING DATE: 1999-11-17
   NUMBER OF SEQ ID NOS: 116
   SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
   LENGTH: 374
   TYPE: PRT
   ORGANISM: Artificial Sequence
   FEATURE:
   OTHER INFORMATION: Description of Artificial Sequence: G protein
   OTHER INFORMATION: chimera
US-09-442-349A-26

  Query Match             56.9%;  Score 1058;  DB 4;  Length 374;
  Best Local Similarity   57.1%;  Pred. No. 5e-99;
  Matches   206;  Conservative   56;  Mismatches   87;  Indels   12;  Gaps    2;

Qy    10  CLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD   69
          ||:::||| ||:  |::|  |:: ||:|||:|||| |:|||||||||||||||||||| ||:
Db    13  CLTEDEKAAARVDQEINRILLEQKKQDRGELKLILLGPGESGKSTFIKQMRIIHGAGYSE   72

Qy    70  EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV  129
          |::|||  |||||||.:|:|||:|||::|:|:|: |::.::|:|| ::| |||||||||||
Db    73  EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA  132

Qy   130  DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI  189
          :|:. || .| || |||||: ||||||||||||||||||||||||||||||||||||||||
Db   133  AAMQWLWRDAGIRACYERRREFHLLDSAVYYLSHLERITEEGYVPTAQDVLRSRMPTTGI  192

Qy   190  IEYPFDLQSVIFRMVDVGGQQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME  249
          :|| || ::: ||||||||||:||||||||||||||:::: :|||||||.|| |::|||||
Db   193  NEYCFSVQKTNLRIVDVGGQQKSERRKWIHCFENVIALILASLSEYDQCLEENNQENRMK  252

Qy   250  ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI  309
          |||||| || | |||:::||||||||.|:|||||| |||| |||::|||:||| ||:||
Db   253  ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPKQDAEAAKRFI  312

Qy   310  LKMFVDL------NPDSD------KIIYSHFTCATDTENIREVFAAVKDTILQLNLKEYN  357
          | |:   |||||   |:: ||||||:||||||||||||| ||| ||:| |::|:||| :|
Db   313  LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTCNIQFVFNAVTDVIIQNNLKYIG  372

Qy   358  L  358
          |
Db   373  L  373
```

```
RESULT 14
US-09-442-349A-26
; Sequence 26, Application US/0944349A

  Query Match             56.9%;  Score 1058;  DB 4;  Length 374;
```

Best Local Similarity 57.1%;  Pred. No. 5e-99;
Matches 206; Conservative 56; Mismatches 87; Indels 12; Gaps 2;

```
Qy    10  CLSEEAKEARRINDEIERHVRRDKRDAARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD  69
          ||:|:|   |  |:||: ||| : |   :  |||||||||||||||||||||||||:|:||:
Db    13  CLTEDEKAAARVDQEINRILLEQKKQDRGELKLLLLGPGESGKSTFIKQMRIIHGAGYSE  72

Qy    70  EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV  129
          ::::|  |||||||  | :|   |  |||   | | :|||    ||    ||    |  ||
Db    73  EERKGFRPLVYQNIFVSMRAMIEAMERLQIPFSRPESKHHASLVMSQDPYKVTTFEKRYA  132

Qy   130  DAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADDPAYLPTQQDVLRVPTTGI  189
          :|: ||: |  ||  | |:|  |: | |  | |::| ||| || |||||||||||||
Db   133  AAMQWLWRDAGIRACYERRREFHLLDSAVYYLSHLERITEEGYVPPTAQDVLRSRMPTTGI  192

Qy   190  IEYPFDLQSVIFRMVDVGGQQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME  249
          :|:|    | || |||||||||||||||||||  | : :|||| ::| ||||: |  ||:
Db   193  NEYCFSVQKINLRIVDVGGQKSERKKWIHCFENVIALIYLASLSEYDQCLEENNQENRMK  252

Qy   250  ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQQRDAQAAREFI  309
          || ||| || ||||:|| :::|:|:|||| | ||:|   :::||: |: || :  : ||
Db   253  ESLALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPSFQGPKQDAEAAKRFI  312

Qy   310  LKMFVDL------NPDSD------KIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYN  357
          |:|:            :|        |:||||| |:|  || : :|  : | :|:|:
Db   313  LDMYTRMYTGCVDGPEGSNRNKETKEIYSHFTCATDTCNIQFVFQAVTDVIIQNNLKYIG  372

Qy   358  L  358
          |
Db   373  L  373
```

Search completed: June 30, 2003, 16:18:45
Job time : 16 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: June 30, 2003, 16:16:21 ; Search time 24 Seconds
(without alignments)
1640.134 Million cell updates/sec

Title: US-09-899-295-2
Perfect score: 1860
Sequence: 1 MTLESIMACCLSEEAKEARR.........VFAAVKDTILQLNLKEVNLV 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 10946833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1860 | 100.0 | 359 | 10 | US-09-899-295-2 | Sequence 2, Appli |
| 2 | 1846 | 99.2 | 359 | 10 | US-09-984-292-1 | Sequence 1, Appli |
| 3 | 1846 | 99.2 | 359 | 10 | US-09-989-497-1 | Sequence 1, Appli |
| 4 | 1842 | 99.0 | 359 | 10 | US-09-984-292-37 | Sequence 37, Appl |
| 5 | 1842 | 99.0 | 359 | 10 | US-09-989-497-37 | Sequence 37, Appl |
| 6 | 1837 | 98.8 | 359 | 10 | US-09-984-292-15 | Sequence 15, Appl |
| 7 | 1837 | 98.8 | 359 | 10 | US-09-989-497-15 | Sequence 15, Appl |
| 8 | 1822 | 98.0 | 359 | 10 | US-09-984-292-9 | Sequence 9, Appli |
| 9 | 1822 | 98.0 | 359 | 10 | US-09-989-497-9 | Sequence 9, Appli |
| 10 | 1819 | 97.8 | 353 | 10 | US-09-984-292-2 | Sequence 2, Appli |
| 11 | 1819 | 97.8 | 353 | 10 | US-09-989-497-2 | Sequence 2, Appli |
| 12 | 1818 | 97.7 | 359 | 10 | US-09-984-292-21 | Sequence 21, Appl |
| 13 | 1818 | 97.7 | 359 | 10 | US-09-989-497-21 | Sequence 21, Appl |
| 14 | 1817 | 97.7 | 353 | 9 | US-09-952-680A-27 | Sequence 27, Appl |
| 15 | 1817 | 97.7 | 353 | 10 | US-09-984-292-16 | Sequence 16, Appl |
| 16 | 1817 | 97.6 | 353 | 10 | US-09-989-497-16 | Sequence 16, Appl |
| 17 | 1815 | 97.6 | 359 | 10 | US-09-984-292-11 | Sequence 11, Appl |
| 18 | 1815 | 97.6 | 359 | 10 | US-09-989-497-11 | Sequence 11, Appl |
| 19 | 1814 | 97.5 | 353 | 10 | US-09-899-295-6 | Sequence 6, Appli |
| 20 | 1814 | 97.5 | 353 | 10 | US-09-899-295-8 | Sequence 8, Appli |
| 21 | 1813 | 97.5 | 359 | 10 | US-09-984-292-26 | Sequence 26, Appl |
| 22 | 1813 | 97.5 | 359 | 10 | US-09-989-497-26 | Sequence 26, Appl |
| 23 | 1811 | 97.4 | 359 | 10 | US-09-984-292-3 | Sequence 3, Appli |
| 24 | 1811 | 97.4 | 359 | 10 | US-09-989-497-3 | Sequence 3, Appli |
| 25 | 1810 | 97.3 | 353 | 10 | US-09-899-295-4 | Sequence 4, Appli |
| 26 | 1807 | 97.2 | 359 | 10 | US-09-984-292-20 | Sequence 20, Appl |
| 27 | 1807 | 97.2 | 359 | 10 | US-09-989-497-20 | Sequence 20, Appl |
| 28 | 1806 | 97.1 | 359 | 10 | US-09-984-292-25 | Sequence 25, Appl |
| 29 | 1806 | 97.1 | 359 | 10 | US-09-989-497-25 | Sequence 25, Appl |
| 30 | 1804 | 97.0 | 359 | 10 | US-09-984-292-10 | Sequence 10, Appl |
| 31 | 1804 | 97.0 | 359 | 10 | US-09-989-497-10 | Sequence 10, Appl |
| 32 | 1803 | 96.9 | 359 | 10 | US-09-984-292-22 | Sequence 22, Appl |
| 33 | 1803 | 96.9 | 359 | 10 | US-09-989-497-22 | Sequence 22, Appl |
| 34 | 1800 | 96.8 | 359 | 10 | US-09-984-292-12 | Sequence 12, Appl |
| 35 | 1800 | 96.8 | 359 | 10 | US-09-989-497-12 | Sequence 12, Appl |
| 36 | 1795 | 96.5 | 353 | 10 | US-09-984-292-8 | Sequence 8, Appli |
| 37 | 1795 | 96.5 | 359 | 10 | US-09-989-497-8 | Sequence 8, Appli |
| 38 | 1793 | 96.4 | 359 | 10 | US-09-984-292-13 | Sequence 13, Appl |
| 39 | 1793 | 96.3 | 359 | 10 | US-09-989-497-13 | Sequence 13, Appl |
| 40 | 1791 | 96.3 | 353 | 10 | US-09-984-292-17 | Sequence 17, Appl |
| 41 | 1791 | 96.3 | 353 | 10 | US-09-989-497-17 | Sequence 17, Appl |
| 42 | 1789 | 96.2 | 359 | 10 | US-09-984-292-23 | Sequence 23, Appl |
| 43 | 1789 | 96.2 | 359 | 10 | US-09-989-497-23 | Sequence 23, Appl |
| 44 | 1784 | 95.9 | 353 | 10 | US-09-984-292-4 | Sequence 4, Appli |
| 45 | 1784 | 95.9 | 353 | 10 | US-09-899-295-5 | Sequence 5, Appli |

ALIGNMENTS

US-09-899-295-2.rapb

RESULT 1
US-09-899-295-2

; Sequence 2, Application US/09899295
; Patent No. US20020127601A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma Deutschland GmbH
; TITLE OF INVENTION: Process for identifying modulators of G protein coupled
; TITLE OF INVENTION: receptors
; FILE REFERENCE: AVE D-2000/A033 englisch
; CURRENT APPLICATION NUMBER: US/09/899,295
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-899-295-2

Query Match     100.0%;  Score 1860;  DB 10;  Length 359;
Best Local Similarity  100.0%;  Pred. No. 1.8e-149;
Matches 359;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```
Qy   1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy 121 VSAFENPVVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 VSAFENPVVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy 181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy 241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
```

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 2
US-09-984-292-1
; Sequence 1, Application US/09984292
; Patent No. US20020128433A1
; GENERAL INFORMATION:
;  APPLICANT: YAO, YONG
;  APPLICANT: XU, HONG
;  TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
;  TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
;  TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
;  FILE REFERENCE: 078003-0280649
;  CURRENT APPLICATION NUMBER: US/09/984,292
;  CURRENT FILING DATE: 2001-10-29
;  PRIOR APPLICATION NUMBER: 60/243,770
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
;  LENGTH: 359
;  TYPE: PRT
;  ORGANISM: Mus sp.
US-09-984-292-1

  Query Match             99.2%;  Score 1846;  DB 10;  Length 359;
  Best Local Similarity   99.2%;  Pred. No. 2.7e-148;
  Matches  356;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 3
US-09-989-497-1
; Sequence 1, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
;  APPLICANT: YAO, YONG
;  APPLICANT: XU, HONG
;  TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
;  TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
;  TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
;  FILE REFERENCE: 078003-0280735
;  CURRENT APPLICATION NUMBER: US/09/989,497
;  CURRENT FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 09/984,292
;  PRIOR FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1

RESULT 4
US-09-984-292-37
; Sequence 37, Application US/09984292
; Patent No. US20020128433A1
; GENERAL INFORMATION:
;  APPLICANT: YAO, YONG
;  APPLICANT: XU, HONG
;  TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
;  TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
;  TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
;  FILE REFERENCE: 078003-0280649
;  CURRENT APPLICATION NUMBER: US/09/984,292
;  CURRENT FILING DATE: 2001-10-29
;  PRIOR APPLICATION NUMBER: 60/243,770
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
;  LENGTH: 359
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-984-292-37

  Query Match             99.0%;  Score 1842;  DB 10;  Length 359;
  Best Local Similarity   99.2%;  Pred. No. 5.9e-148;
  Matches  356;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
         |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180

Qy   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKDLLEEKIMYSHLVDYFPEYDGPQR 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

**Left column:**

```
RESULT 5
US-09-989-497-37
; Sequence 37, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-989-497-37

Query Match          99.0%;  Score 1842;  DB 10;  Length 359;
Best Local Similarity  99.2%;  Pred. No. 5.9e-148;
Matches  356;  Conservative  1;  Mismatches  2;  Indels  0;  Gaps  0;

Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy     61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNLV 359
            ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    301 DAQAAREFILKMFVDLNPDSDKINYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNLV 359
```

**Right column:**

```
US-09-984-292-15
; Sequence 15, Application US/09984292
; Patent No. US20020128433A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-292-15

Query Match          98.8%;  Score 1837;  DB 10;  Length 359;
Best Local Similarity  98.9%;  Pred. No. 1.6e-147;
Matches  355;  Conservative  1;  Mismatches  3;  Indels  0;  Gaps  0;

Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
            |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db      1 MTLESIMACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy     61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IIHGGSYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNLV 359
            ||||||||||||||||||||||| ||||||||||||||||||||||||| |||||||| |
Db    301 DAQAAREFILKMFVDLNPDSDKINYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNAV 359

RESULT 7
US-09-989-497-15
; Sequence 15, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 359
; TYPE: PRT
```

RESULT 6

## Left column

```
; ORGANISM: Homo sapiens
US-09-989-497-15

Query Match           98.8%;   Score 1837;  DB 10;  Length 359;
Best Local Similarity 98.9%;   Pred. No. 1.6e-147;
Matches 355;  Conservative 1;  Mismatches 3;  Indels 0;  Gaps 0;

Qy    1 MTLESIMACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMR 60
         ||||||||||||||||||||||||||||:|::||||:||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERQLRDKRDAARRELKLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db  301 DAQAAREFILKMFVDLNPDSDKINYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV 359
```

```
RESULT 8
US-09-984-292-9
; Sequence 9, Application US/0984292
; Patent No. US2002128433A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09/984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-984-292-9

Query Match           98.0%;   Score 1822;  DB 10;  Length 359;
Best Local Similarity 98.3%;   Pred. No. 2.9e-146;
Matches 353;  Conservative 2;  Mismatches 4;  Indels 0;  Gaps 0;

Qy    1 MTLESIMACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMR 60
         |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||:::|::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFDVPDYAAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

## Right column

```
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

```
RESULT 9
US-09-989-497-9
; Sequence 9, Application US/09899497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09/989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; SOFTWARE: Patentin Ver. 2.1
; NUMBER OF SEQ ID NOS: 42
; SEQ ID NO 9
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-9

Query Match           98.0%;   Score 1822;  DB 10;  Length 359;
Best Local Similarity 98.3%;   Pred. No. 2.9e-146;
Matches 353;  Conservative 2;  Mismatches 4;  Indels 0;  Gaps 0;

Qy    1 MTLESIMACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMR 60
         |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||:::|::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 VSAFDVPDYAAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180

Qy  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

```
RESULT 10
US-09-984-292-2
; Sequence 2, Application US/0984292
; Patent No. US2002128433A1
; GENERAL INFORMATION:
; APPLICANT: YAO, YONG
; APPLICANT: XU, HONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
```

RESULT 11
US-09-989-497-2
; Sequence 2, Application US/09989497
; Patent No. US20020143151A1
; GENERAL INFORMATION:
; APPLICANT: XU, HONG
; APPLICANT: YAO, YONG
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; FILE REFERENCE: 078003-0280735
; CURRENT APPLICATION NUMBER: US/09,989,497
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/984,292
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-989-497-2

Query Match             97.8%;  Score 1819;  DB 10;  Length 353;
Best Local Similarity   99.2%;  Pred. No. 5.1e-146;
  Matches  350;  Conservative   2;  Mismatches   1;  Indels   0;  Gaps   0;

Qy          7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLLGTGESGKSTFIKQMRIIHGSG 66
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 MACCLSEEAKEARRINDEIERQLRDKRDARRELKLILLLGTGESGKSTFIKQMRIIHGSG 60

Qy         67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 120

Qy        127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 180

Qy        187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 240

Qy        247 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 300

Qy        307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353


RESULT 12
US-09-984-292-21
; Sequence 21, Application US/09984292
; Patent No. US20020128433A1
; GENERAL INFORMATION:
; APPLICANT: XU, HONG
; APPLICANT: YAO, YONG
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-0280649
; CURRENT APPLICATION NUMBER: US/09,984,292
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-292-21

Query Match             97.7%;  Score 1818;  DB 10;  Length 359;
Best Local Similarity   98.3%;  Pred. No. 6.3e-146;
  Matches  353;  Conservative   1;  Mismatches   5;  Indels   0;  Gaps   0;

Qy          1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLLGTGESGKSTFIKQMR 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLLGTGESGKSTFIKQMR 60

Qy         61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IIHGSDYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy        121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
              |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VSAFDVPDYVAAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy        181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy        241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy        301 DAQAAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

RESULT 13
US-09-989-497-21
Sequence 21, Application US/09989497
Patent No. US2002014315lA1
GENERAL INFORMATION:
APPLICANT: XU, HONG
APPLICANT: YAO, YONG
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
FILE REFERENCE: 078003-0280735
CURRENT APPLICATION NUMBER: US/09/989,497
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 09/984,292
PRIOR FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 60/243,770
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 21
LENGTH: 359
TYPE: PRT
ORGANISM: Homo sapiens
US-09-989-497-21

Query Match             97.7%;  Score 1818;  DB 10;  Length 359;
Best Local Similarity   98.3%;  Pred. No. 6.3e-146;
Matches  353;  Conservative   1;  Mismatches   5;   Indels   0;   Gaps   0;

Qy       1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy      61 IIHGSDYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
           |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 IIHGSDYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy     121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
           ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VSAFDVPDYAAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180

Qy     181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy     241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKDLLEEKIMYSHLVDYFPEYDGPQR 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy     301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNLV 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNLV 359


RESULT 14
US-09-952-680A-27
Sequence 27, Application US/09952680A
Publication No. US20030087239A1
GENERAL INFORMATION:
APPLICANT: Stanton, Marty
APPLICANT: Epstein, David
APPLICANT: Hamaguchi, No. US20030087239A1uko
TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same
FILE REFERENCE: 23239-501
CURRENT APPLICATION NUMBER: US/09/952,680A
CURRENT FILING DATE: 2001-09-13
PRIOR APPLICATION NUMBER: 60/232,454
PRIOR FILING DATE: 2000-09-13
NUMBER OF SEQ ID NOS: 75

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 27
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-952-680A-27

Query Match             97.7%;  Score 1817;  DB 9;  Length 353;
Best Local Similarity   99.2%;  Pred. No. 7.5e-146;
Matches  350;  Conservative   1;  Mismatches   2;   Indels   0;   Gaps   0;

Qy       7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MACCLSEEAKEARRINDEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 60

Qy      67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 120

Qy     127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 186
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 180

Qy     187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 246
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 240

Qy     247 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 306
           ||||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||
Db     241 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 300

Qy     307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNAV 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLKEYNAV 353


RESULT 15
US-09-984-292-16
Sequence 16, Application US/09984292
Patent No. US2002012843lA1
GENERAL INFORMATION:
APPLICANT: XU, HONG
APPLICANT: YAO, YONG
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
FILE REFERENCE: 078003-0280649
CURRENT APPLICATION NUMBER: US/09/984,292
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 60/243,770
PRIOR FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 16
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-984-292-16

```
Db   121   PYVDAIKSLWNDPGIQEBCYDRREYQLSDSTKYLNDLDRVADPAYLPTQQDVLRVRVPT  180

Qy   187   TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181   TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy   247   RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241   RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  300

Qy   307   EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
           |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   301   EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNAV  353
```

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:               June 30, 2003, 16:14:36 ; Search time 19 Seconds
                                            (without alignments)
                                            1816.435 Million cell updates/sec

Title:                US-09-899-295-2
Perfect score:        1860
Sequence:             1 MTLESIMACCLSEEAKEARR..........VFAAVKDTIIQLNLKEYNLV 359

Scoring table:        BLOSUM62
                      Gapop 10.0 , Gapext 0.5

Searched:             283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:         283224

Post-processing:      Minimum Match 0%
Maximum DB seq length: 2000000000
Minimum DB seq length: 0

Database :            PIR_73:*
                      1: pir1:*
                      2: pir2:*
                      3: pir3:*
                      4: pir4:*

                      Maximum Match 100%
                      Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1857 | 99.8 | 359 | 1 | RGMSQ | GTP-binding regula |
| 2 | 1844 | 99.1 | 359 | 2 | S71963 | GTP-binding protei |
| 3 | 1799 | 96.7 | 359 | 2 | RGHYO2 | GTP-binding regula |
| 4 | 1709 | 91.9 | 359 | 2 | S45699 | GTP-binding regula |
| 5 | 1685 | 90.6 | 359 | 2 | S30359 | GTP-binding regula |
| 6 | 1682 | 90.4 | 359 | 1 | S45700 | G-alpha-11 protein |
| 7 | 1681 | 90.4 | 359 | 1 | RGMS1 | GTP-binding regula |
| 7 | 1681 | 90.4 | 359 | 1 | RGHUGV | GTP-binding regula |
| 8 | 1667 | 89.6 | 353 | 2 | B40891 | GTP-binding protei |
| 9 | 1538 | 82.7 | 355 | 2 | A40891 | GTP-binding protei |
| 10 | 1535 | 82.5 | 355 | 2 | A41534 | GTP-binding protei |
| 11 | 1524 | 81.9 | 355 | 2 | S34347 | GTP-binding regula |
| 12 | 1513 | 81.3 | 355 | 2 | T15288 | hypothetical prote |
| 13 | 1469.5 | 79.0 | 354 | 2 | S33309 | GTP-binding regula |
| 14 | 1425.5 | 76.6 | 360 | 2 | JN0115 | GTP-binding regula |
| 15 | 1066 | 57.3 | 374 | 2 | B41534 | GTP-binding protei |
| 16 | 1044 | 56.1 | 374 | 2 | A41096 | GTP-binding protei |
| 17 | 913 | 49.1 | 353 | 2 | S71965 | GTP-binding regula |
| 18 | 904 | 48.6 | 198 | 2 | S71964 | GTP-binding protei |
| 19 | 898.5 | 48.3 | 354 | 1 | RGHU01 | GTP-binding regula |
| 20 | 898.5 | 48.3 | 354 | 1 | RGB01I | GTP-binding regula |
| 21 | 897.5 | 48.3 | 354 | 1 | RGRT1I | GTP-binding regula |
| 22 | 895.5 | 48.1 | 354 | 2 | S40508 | GTP-binding regula |
| 23 | 893 | 48.0 | 353 | 2 | JC7661 | G protein alpha su |
| 24 | 890 | 47.8 | 353 | 2 | T50482 | G protein alpha ch |
| 25 | 888.5 | 47.8 | 354 | 1 | RGXL1I | G protein alpha ch |
| 26 | 888.5 | 47.8 | 354 | 2 | I50237 | GTP-binding regula |
| 27 | 887.5 | 47.7 | 354 | 2 | S40509 | GTP-binding regula |
| 28 | 886.5 | 47.7 | 354 | 1 | RGHUO2 | G-protein - chicke |
| 29 | 886.5 | 47.7 | 354 | 2 | S28157 | GTP-binding regula |
| 30 | 883 | 47.5 | 355 | 2 | S28158 | GTP-binding regula |
| 31 | 882.5 | 47.4 | 354 | 1 | RGRT13 | GTP-binding regula |
| 32 | 882.5 | 47.4 | 354 | 1 | RGHYO2 | GTP-binding regula |
| 33 | 882 | 47.4 | 355 | 2 | I50238 | Gi2 protein alpha- |
| 34 | 881.5 | 47.4 | 354 | 1 | RGHUI3 | GTP-binding regula |
| 35 | 878.5 | 47.2 | 354 | 2 | S27014 | GTP-binding regula |
| 36 | 877.5 | 47.2 | 354 | 2 | S24362 | GTP-binding regula |
| 37 | 877 | 47.2 | 355 | 1 | RGHUI2 | GTP-binding regula |
| 38 | 877 | 47.0 | 355 | 2 | A61031 | GTP-binding regula |
| 39 | 876.5 | 47.0 | 354 | 2 | S28159 | GTP-binding regula |
| 40 | 873.5 | 46.9 | 354 | 2 | RGMSO2 | GTP-binding regula |
| 41 | 872.5 | 46.9 | 354 | 2 | S27013 | GTP-binding regula |
| 42 | 872.5 | 46.9 | 354 | 2 | RGRTO2 | GTP-binding regula |
| 43 | 871.5 | 46.9 | 354 | 1 | RGRTI2 | GTP-binding regula |
| 44 | 871 | 46.8 | 355 | 1 | RGRT12 | GTP-binding regula |
| 45 | 865.5 | 46.5 | 354 | 1 | RGFFO1 | GTP-binding regula |

## ALIGNMENTS

### RESULT 1

RGMSQ

GTP-binding regulatory protein Gq alpha chain - mouse

GTP-binding regulatory protein Gq alpha chain; heterotrimeric guanine nucleotide binding protein Gq alpha chain
N;Alternate names: guanine nucleotide binding protein Gq alpha chain; heterotrimeric
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A38414
R;Strathmann, M.; Simon, M.I.
A;Residues: 1-359 <STR>
A;Cross-references: GB:M55412; NID:g193501; PIDN:AAA63306.1; PID:g193502
A;Title: G protein diversity: a distinct class of alpha subunits is present in verteb
A;Reference number: A38414; MUID:91067657; PMID:2123549
Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990
A;Accession: A38414
A;Molecule type: mRNA
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that re
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
ains. The beta and gamma chains, required for GTPase activity, appear to be common to
rase; it is specific for each type of G protein.
C;Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transductio
F;46-53/Region: GTP binding; nucleotide-binding motif A (P-loop)
F;274-277/Region: GTP-binding NKXD motif
F;52/Binding site: GTP (Lys) #status predicted
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match      99.8%;    Score 1857;    DB 1;    Length 359;
Best Local Similarity      99.7%;    Pred. No. 1.4e-126;
Matches   358;    Conservative     1;    Mismatches    0;    Indels    0;    Gaps    0;

Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy     61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120

Qy    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPSYLPTQQDVL 180

Qy    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV 359

RESULT 2
S71963
GTP-binding protein alpha-q - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 02-Feb-2001
C;Accession: S71963
R;Johnson, G.J.; Leis, L.A.; Dunlop, P.C.
Biochem. J. 318, 1023-1031, 1996
A;Title: Specificity of G-alpha(q) and G-alpha(11) gene expression in platelets and eryt
A;Reference number: S71963; MUID:96433124; PMID:8836152
A;Accession: S71963
A;Molecule type: mRNA
A;Residues: 1-359 <JOH>
A;Cross-references: EMBL:L76256; NID:g1478071; PIDN:AAB39498.1; PID:g1478072
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction
F;46-53/Region: nucleotide-binding motif A (p-loop)
F;274-277/Region: GTP-binding NKXD motif

Query Match                  99.1%;  Score 1844;  DB 2;  Length 359;
Best Local Similarity        99.2%;  Pred. No. 1.2e-125;
Matches    356;  Conservative   1;  Mismatches   2;  Indels    0;  Gaps    0;

Qy      1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR  60
           ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db      1  MTLESIMACCLSEEAKEARRINDEIEQLRDKRDARRELKLLLLGTGESGKSTFIKQMR  60

Qy     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120

Qy    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180

Qy    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240

Qy    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300

Qy    301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359

─────────────────────────────────────────────────

RESULT 3
S45699
GTP-binding regulatory protein alpha chain q - African clawed frog
N;Alternate names: G-alpha-q protein
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C;Accession: S45699
R;Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battey, J.F.
FEBS Lett. 348, 89-92, 1994
A;Title: Neuromedin B receptor, expressed in Xenopus laevis oocytes, selectively couples
A;Reference number: S45699; MUID:94298961; PMID:8026589
A;Accession: S45699
A;Molecule type: mRNA
A;Residues: 1-359 <SHA>
A;Cross-references: GB:U10502; NID:g505689; PIDN:AAA52189.1; PID:g505690
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction
F;46-53/Region: nucleotide-binding motif A (p-loop)
F;274-277/Region: GTP-binding NKXD motif
F;52/Binding site: GTP (Lys) #status predicted
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match                  96.7%;  Score 1799;  DB 2;  Length 359;

─────────────────────────────────────────────────

Best Local Similarity        96.4%;  Pred. No. 2.1e-122;
Matches    346;  Conservative   6;  Mismatches   7;  Indels    0;  Gaps    0;

Qy      1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR  60
           ||||||||||||||||||||||||||||||:||:||||||||||||||||||||||||||
Db      1  MTLESIMACCLSEEAKEARRINDEIEQLRDKRDARRELKLLLLGTGESGKSTFIKQMR  60

Qy     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120

Qy    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180

Qy    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240

Qy    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300

Qy    301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359

─────────────────────────────────────────────────

RESULT 4
S30359
GTP-binding regulatory protein G alpha chain, phospholipase C-activating - turkey
N;Alternate names: phospholipase C-activating G protein
C;Species: Meleagris gallopavo (common turkey)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Feb-2001
C;Accession: S30359; S30360
R;Maurice, D.H.; Waldo, G.L.; Morris, A.J.; Nicholas, R.A.; Harden, T.K.
Biochem. J. 290, 765-770, 1993
A;Title: Identification of Galpha(11) as the phospholipase C-activating G-protein of
A;Reference number: S30359; MUID:93207527; PMID:8457205
A;Accession: S30359
A;Molecule type: mRNA
A;Residues: 1-359 <MAU>
A;Cross-references: GB:X73072; NID:g312254; PIDN:CAA51530.1; PID:g312255
A;Experimental source: blood
A;Accession: S30360
A;Molecule type: protein
A;Residues: 78-92;121-132;158-180;253-256;307-312;339-345;355-359 <MAW>
A;Experimental source: erythrocytes
A;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; P-loop
F;46-53/Region: nucleotide-binding motif A (P-loop)
F;156-158/Region: GTP-binding SAK/L motif
F;274-277/Region: GTP-binding NKXD motif
F;52/Binding site: GTP (Lys) #status predicted
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match                  91.9%;  Score 1709;  DB 2;  Length 359;
Best Local Similarity        90.3%;  Pred. No. 6.4e-116;
Matches    324;  Conservative  22;  Mismatches  13;  Indels    0;  Gaps    0;

Qy      1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR  60
           ||||||:|||||||||||:|||||| || :||:||||||||||||||||||||||||||
Db      1  MTLESMACCLSDEVKESKRINAEIEKQLRDKRDARRELKLLLLGTGESGKSTFIKQMR  60

Qy     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
           ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQSMIRAMETLKILYKYEQNKANAVLVREVDVEK  120

Qy    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
           | || |||||:|||||||||||||||||||||||||||:||||||||||||||||||||
Db    121  VWTFEQPYVSAIKTLWNDPGIQECYDRRREYQLSDSAKYYLSDVDRIATPGYLPTQQDVL  180

Qy    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240

## RESULT 5

S45700

G-alpha-11 protein - African clawed frog

```
A;Molecule type: mRNA
A;Accession: S45700
A;Reference number: S45699; MUID:94289961; PMID:8026589
A;Title: Neuromedin B receptor, expressed in Xenopus laevis oocytes, selectively couples
R;Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battey, J.F.
FEBS Lett. 348, 89-92, 1994
A;Accession: S45700
C;Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C;Species: Xenopus laevis (African clawed frog)
C;Cross-references: GB:U10494; NID:g505687; PIDN:AAA52188.1; PID:g505688
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; P-loop
F;46-53/Region: nucleotide-binding motif A (P-loop)
F;274-277/Region: GTP-binding NKXD motif
```

```
Query Match     90.6%;     Score 1685;     DB 2;     Length 359;
Best Local Similarity  88.9%;
Matches 319;    Conservative  25;   Mismatches  15;   Indels   0;   Gaps   0;

Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLGTGESGKSTFIKQMR  60
         ||||||||||:||:||||||||||:|::||:|:|||| |||||||||||||||||||||
Db     1 MTLDSTMACCLSEEVVEKSKRINAEIEKQLRDKKDSRRELKLLLGTGESGKSTFIKQMR  60

Qy    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
         ||||||||||||||||||||||||||||||||||||||:::||||:||||||||||||||
Db    61 IIHGSGYSEEDKKGFTKLVFQNIFTAMQSMIRAMETLKILYKYEQNKANAQVREVDVEK  120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
         ||::|||||||||||:||||||||||||||||||||||||||||:|:||:|||||||||||
Db   121 VCTFEQPYVNAIKNLWSDPGIQECYDRRREYQLSDSTKYYLTDVDRISKPGYLPTQQDVL 180

Qy   181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
         ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db   181 RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

## RESULT 6

RGMS1

GTP-binding regulatory protein G11 alpha chain - mouse

```
A;Molecule type: mRNA
A;Accession: B38414
A;Reference number: A38414; MUID:91067657; PMID:2123549
A;Title: G protein diversity: a distinct class of alpha subunits is present in vertebra
R;Strathmann, M.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990
A;Accession: B38414
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Species: Mus musculus (house mouse)
C;Alternate names: guanine nucleotide binding protein G11 alpha chain; heterotrimeric G-
N;Alternate names: guanine nucleotide binding protein G11 alpha chain; heterotrimeric G
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
```

```
Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLGTGESGKSTFIKQMR  60
Db     1 MTLDSTMACCLSEEVVEKSKRINAEIEKQLRDKKDSRRELKLLLGTGESGKSTFIKQMR  60

Qy    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
Db    61 IIHGSGYSEEDKKGFTKLVFQNIFTAMQSMIRAMETLKILYKYEQNKANAQVREVDVEK  120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
Db   121 VTTFEHQYVNAIKLWSDPGIQECYDRRREFQLSDSAAKYLTDVDRIATWGYLPTQQDVL 180

Qy   181 RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
Db   181 RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEDKIMYSHLVDYFPEFDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
Db   301 DAATAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQHNLKEYNLV 359
```

```
Query Match     90.4%;     Score 1681;     DB 1;     Length 359;
Best Local Similarity  89.1%;
Matches 320;    Conservative  23;   Mismatches  16;   Indels   0;   Gaps   0;
```

## RESULT 7

RGHUGY

GTP-binding regulatory protein Gy alpha chain - human

```
A;Molecule type: mRNA
A;Accession: A39394
A;Reference number: A39394; MUID:91219481; PMID:1902575
A;Title: Guanine nucleotide-binding regulatory proteins in retinal pigment epithelial
R;Jiang, M.; Pandey, S.; Tran, V.T.; Fong, H.K.W.
Proc. Natl. Acad. Sci. U.S.A. 88, 3907-3911, 1991
A;Accession: A39394
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Species: Homo sapiens (man)
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Map position: 19p13.3-19p13.3
A;Cross-references: GDB:132587; OMIM:139313
A;Gene: GDB:GNA11
A;Cross-references: GDB:GNA11
C;Cross-references: GB:M69013; NID:g1836690; PIDN:AAA88624.1; PID:g1836691
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that re
ans. The beta and gamma chains, required for GTPase activity, appear to be common to
rase; it is specific for each type of G protein.
C;Genetics:
```

```
Qy     1 MTLESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLGTGESGKSTFIKQMR  60
         ||||:|.||.|:|||||||||||:||:|:||:|.|:|||||||||||||||||||||||
Db     1 MTLESMMACCLSDEVKESKRINAEIEKQLRDKRDARRELKLLLGTGESGKSTFIKQMR  60

Qy    61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
Db    61 IIHGAGYSEEDKRGFTKLVYQNIFTAMQAMVRAMETLKILYKYEQNKANALLIREVDVEK 120

Qy   121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
Db   121 VTTFEHQYVNAIKTLWLWSDPGVQECYDRRREFQLSDSAKYLTDVDRIATVGYLPTQQDVL 180

Qy   181 RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
Db   181 RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
Db   241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKILHSHLVDYFPEFDGPQR 300

Qy   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
Db   301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

```
A;Residues: 1-359 <STR>
A;Cross-references: GB:M55411; NID:g193499; PIDN:AAA63305.1; PID:g193500
C;Comment: The G proteins are a family of guanine nucleotide-binding proteins that re
ains. It is specific for each type of G protein.
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transductio
F;46-53/Region: nucleotide-binding motif A (P-loop)
F;274-277/Region: GTP-binding NKXD motif
F;52/Binding site: GTP (Lys) #status predicted
F;183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match     90.4%;     Score 1682;     DB 1;     Length 359;
Best Local Similarity  88.6%;
Matches 318;    Conservative  28;   Mismatches  13;   Indels   0;   Gaps   0;
```

Qy      1  MTLESTMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMR     60
           |||||::|||||::|::|:||||:||::||:||||||||||||||||||||||||||||
Db      1  MTLESMMACCLSDEVKESKRINAEIEKQLRDKRDARRELKLLLLGTGESGKSTFIKQMR     60

Qy     61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK    120
           |||:||||:|||||||||||||||||||||||||||::|||:|||||:||:||||||||
Db     61  IIHGAGYSEEDKRGFTKLVYQNIFTAMQAMIRAMEFLKILYKYEQNKANALLIREVDVEK    120

Qy    121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVL    180
           |::||:|||:|||:||:|||||||||||||||||||:|||| |:||:|||:||||||||||
Db    121  VTTFEHQYVSAIKTLWEDPGIQECYDRRREYQLSDSAKYILTDVDRIATGYLPTQQDVL    180

Qy    181  RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV    240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV    240

Qy    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR    300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR    300

Qy    301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV    359
           :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  EPQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV    359

RESULT 8
B40891
GTP-binding protein GL2 alpha chain - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001
C;Accession: B40891
R;Nakamura, F.; Ogata, K.; Shiozaki, K.; Kameyama, K.; Ohara, K.; Haga, T.; Nukada, T.
J. Biol. Chem. 266, 12676-12681, 1991
A;Title: Identification of two novel GTP-binding protein alpha-subunits that lack appare
A;Reference number: A40891; MUID:9128G303; PMID:1905731
A;Accession: B40891
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-353 <NAK>
A;Cross-references: GB:D90336; NID:g217567; PIDN:BAA14350.1; PID:g217568
A;Keywords: GTP binding; nucleotide binding; P-loop
F;40-47/Region: nucleotide-binding motif A (P-loop)
F;150-152/Region: GTP-binding SAK/L motif
F;268-271/Region: GTP-binding NKXD motif

Query Match    89.6%;  Score 1667;  DB 2;  Length 353;
Best Local Similarity    89.5%;
Matches 316;  Conservative 23;  Mismatches 14;  Indels 0;  Gaps 0;

Qy      7  MACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG     66
           |||||::|||:||:||:||:||:||||||||||||||||||||||||||||||||||||:||
Db      1  MACCLSDEVKESKRINAEIEKQLRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGAG     60

Qy     67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN    126
           ||:|||||||||||||||||||||||||||::|||:|||||:|||:|||||||||||:||
Db     61  YSEEDKRGFTKLVYQNIFTAMQAMIRAMEFLKILYKYEQNKANALLIREVDVEKVTTFEH    120

Qy    127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVLRVRVPT    186
           :|||:|||:||:||||||||||||||||||:|||||:|:||:|||:|||||||||||||||
Db    121  QYVSAIKTLWEDPGIQECYDRRREYQLSDSAKYILTDVDRIATGYLPTQQDVLRVRVPT    180

Qy    187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN    246
           ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN    240

Qy    247  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR    306
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db    241  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEFDGPQRDAQAAR    300

Qy    307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV        359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV        353

RESULT 9
A40891
GTP-binding protein GL1 alpha chain - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001
C;Accession: A40891
R;Nakamura, F.; Ogata, K.; Shiozaki, K.; Kameyama, K.; Ohara, K.; Haga, T.; Nukada, T
J. Biol. Chem. 266, 12676-12681, 1991
A;Title: Identification of two novel GTP-binding protein alpha-subunits that lack app
A;Reference number: A40891; MUID:9128G303; PMID:1905731
A;Accession: A40891
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-355 <NAK>
A;Cross-references: GB:D90335; NID:g217565; PIDN:BAA14349.1; PID:g217566
A;Superfamily: GTP-binding regulatory protein Gs alpha chain
A;Keywords: GTP binding; nucleotide binding; P-loop
F;42-49/Region: nucleotide-binding motif A (P-loop)
F;152-154/Region: GTP-binding SAK/L motif
F;270-273/Region: GTP-binding NKXD motif

Query Match    82.7%;  Score 1538;  DB 2;  Length 355;
Best Local Similarity    82.3%;
Matches 289;  Conservative 33;  Mismatches 29;  Indels 0;  Gaps 0;

Qy      9  CCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG      66
           ||||:|:||:||||||||:||||||||||||||||||||||||||||||||||||||||
Db      5  CCLSAEEKESQRISAEIERQLRRDKKDARRELKLLLLGTGESGKSTFIKQMRIIHGSYS    64

Qy     69  DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPY    128
           :|||||||||||||||||||||||||:::|:||:|:|:|::||:||||||:||||:||:|
Db     65  DERKGFTKLVYQNIFTAMQAMIRAMDTLKIQYVCEQNKENAQLIREVEVDKVSTLSRDQ    124

Qy    129  VDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVLRVRVPTTG    188
           :|||||||:||||||||||||||||||||:||||:|:|:|:|||:||||||||||||||||
Db    125  VEAIKQLWQDPGIQECYDRRREYQLSDSAKYILTDIDRIAMPAFVPTQQDVLRVRVPTTG    184

Qy    189  IIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM    248
           ||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185  IIEYPFDLENIIFRMVDVGGQRSERRKWIHCFESVSIFLVALSEYDQVLAECDNENRM    244

Qy    249  EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREF    308
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||:|
Db    245  EESKALFKTIITYPWFLNSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF    304

Qy    309  ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIIQLNLKEYNLV          359
           |||||||||||||||||||||||||||||||||||||||:::||||:|:|||
Db    305  ILKLYQDQNPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV          355

RESULT 10
A41534
GTP-binding protein alpha-14 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 02-Feb-2001
C;Accession: A41534; E33833
R;Wilkie, T.M.; Scherle, P.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053, 1991
A;Title: Characterization of G-protein alpha subunits in the G-q class: expression in
A;Reference number: A41534; MUID:92052208; PMID:1946421
A;Accession: A41534
A;Molecule type: mRNA
A;Residues: 1-355 <WIL>
A;Cross-references: GB:M80631; NID:g193568; PIDN:AAA83222.1; PID:g193569
R;Strathmann, M.; Wilkie, T.M.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 86, 7407-7409, 1989
A;Title: Diversity of the G-protein family: sequences from five additional alpha-subu
A;Reference number: A33833; MUID:90017488; PMID:2508088
A;Accession: E33833
A;Molecule type: mRNA

A;Residues: 217-267 <STR>
A;Cross-references: GB:M57616; NID:g193380; PIDN:AAA63304.1; PID:g193381; GB:M26739
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; P-loop
F;42-49/Region: nucleotide-binding motif A (P-loop)
F;152-154/Region: GTP-binding SAK/L motif
F;270-273/Region: GTP-binding NKXD motif

Query Match      82.5%;   Score 1535;   DB 2;   Length 355;
Best Local Similarity   82.1%;   Pred. No. 2.3e-103;
Matches 288;   Conservative 35;   Mismatches 28;   Indels 0;   Gaps 0;

Qy        9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 68
            |||  |:| |||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        5 CCLSAEEKESQRISAEIERHVRRDKKDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 64

Qy       69 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPY 128
            |||||||||||||||||||||||||||:||:|::|:|||||||||||||||||:|:|:|
Db       65 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ 124

Qy      129 VDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVLRVRVPTTG 188
            ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      125 VAAIKQLWLDPGIQECYDRRREYQLSDSAKYLTDIERIAMPSFVPTQQDVLRVRVPTTG 184

Qy      185 IIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
            ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db      245 IIEYPFDLENIIFRMVDVGGQRSERRKWIHCFESVTSIMFLVALSEYDQVLAECDNENRM 244

Qy      189 IIEYPFDLENIIFRMVDVGGQRSERRKWIHCFESVTSIMFLVALSEYDQVLAECDNENRM 248
            (alignment block)

Qy      249 EESKALFRTIIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAREF 308
            |||||||||||||||||||||||||||||||||||||||||||||:|||:||:||||||
Db      245 EESKALFRTIIITPWFLNSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDE 304

Qy      309 ILKMFVDLNPSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      305 ILKLYQDNPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355

RESULT 11
S34347
GTP-binding regulatory protein Gq alpha chain - great pond snail
C;Species: Lymnaea stagnalis (great pond snail)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001
C;Accession: S65461; S34347
R;Knol, J.C.; Ramatsingh, S.; van Kesteren, E.R.; van Minnen, J.; Planta, R.J.; van Hee
Eur. J. Biochem. 230, 193-199, 1995
A;Title: Cloning of a molluscan G protein alpha subunit of the Gq class which is express
A;Reference number: S65461; MUID:95324523; PMID:7601100
A;Accession: S65461
A;Molecule type: mRNA
A;Residues: 1-353 <KNO>
A;Cross-references: EMBL:Z23106; NID:g312629; PIDN:CAA80653.1; PID:g312630
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction
F;40-48/Region: nucleotide-binding motif A (P-loop)
F;150-152/Region: GTP-binding SAK/L motif
F;268-271/Region: GTP-binding NKXD motif
F;46/Binding site: GTP (Lys) #status predicted
F;177/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match      81.9%;   Score 1524;   DB 2;   Length 353;
Best Local Similarity   81.6%;   Pred. No. 1.4e-102;
Matches 288;   Conservative 29;   Mismatches 36;   Indels 0;   Gaps 0;

Qy        7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||:|
Db        1 MACCIPDELKEQKRINQEIERQLKRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGAG 60

Qy       67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
            |||||:|||||||||||||||||||||||||||||||:|
Db       61 YSDEDKRSHIKIVYQNIFMAMHAMIRAMDTLNIQYINPANRENGNMIRQIDYETVTTFDK 120

Qy      127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVLRVRVPT 186
            |  |||||||||||||||||||||||||||:|
Db      121 YSEDEDKRSHIKIVYQNIFMAMHAMIRAMDTLNIQYINPANRENGNMIRQIDYETVTTFDK 120

---

Db      121 PCVDAIISLWNDDGIQECYDRRREYQLTDSAKYILDSVERISQQDYLPFLQDILRVRVPTI 180
            | |||| ||||| |||||||||||||||| |||||| :|:|::||||:|||:||||||
Qy      187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 246
            |||||||||  ||||||||||||||||||||||||||||||||||||||||||:|||||
Db      181 TGIIEYPFDLDSIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 240

Qy      247 RMEESKALFRTIIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAR 306
            |||||||||||||||||||||||||||||||||||||||||||||||||||:| ||||
Db      241 RMEESKALFRTIIITPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEDGPKKEASTAR 300

Qy      307 EFILKMFVDLNPSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 EFILKMFVELNPDPDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353

RESULT 12
T15288
hypothetical protein M01D7.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
C;Accession: T15288
R;Gattung, S.; Goela, D.; Wilson, R.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid M01D7.
A;Reference number: 218322
A;Accession: T15288
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-355 <GAT>
A;Cross-references: EMBL:AF003739; NID:g2105482; PID:g2105489; PIDN:AAB58071.1; GSPDB
A;Experimental source: strain Bristol N2; clone M01D7
C;Genetics:
A;Gene: CESP:M01D7.7
A;Map position: 1
A;Introns: 40/1, 82/3, 155/2, 198/2, 241/3, 293/1, 329/1
C;Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match      81.3%;   Score 1513;   DB 2;   Length 355;
Best Local Similarity   82.0%;   Pred. No. 8.7e-102;
Matches 291;   Conservative 26;   Mismatches 36;   Indels 2;   Gaps 1;

Qy        7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
            |||||||||:|:||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MACCLSEEAREQKRINQEIEKQLQRDKRNARRELKLLLLGTGESGKSTFIKQMRIIHGCG 60

Qy       67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPY--KYEHNKAHAQLVREVDVEKVSAF 124
            |||||||||||||||||||||||||||||||||||  |:||||||||||||||||||||
Db       61 YSEEDKRAHIRLVIQVNVFMAIQSMIRAMDLDIKFGNESSEELQEKAAVVREVDFESVISF 120

Qy      125 ENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVLRVRV 184
            :||||||||||||||||||||||||||||||:|||||| :||||||||||||||||||||
Db      121 EEPVVSYIKELWEDSGIQECYDRRREYQLTDSAKYLSDLRRIAVPDLVLPTEQDILRVRV 180

Qy      185 PTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDN 244
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db      181 PTTGIIEYPFDLEQIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVECDN 240

Qy      245 ENRMEESKALFRTIIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQ 304
            ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ENRMEESKALFRTIIITYPWFNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPPRDPIA 300

Qy      305 AREFILKMFVDLNPSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 355
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AREFILKMFVDLNPDADKIIYSHFTCATDTENIRFVFAAVKDTILQHNLKEYNLV 355

RESULT 13
S33309
GTP-binding regulatory protein Gq alpha chain - northern European squid
C;Species: Loligo forbesi (northern European squid)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
C;Accession: S33309
R;Ryba, N.J.P.; Findlay, J.B.C.; Reid, J.D.

```
Biochem. J. 292, 333-341, 1993
A;Title: The molecular cloning of the squid (Loligo forbesi) visual Gq-alpha subunit and
A;Reference number: S33309; MUID:93277493; PMID:8503868
A;Accession: S33309
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-354 <RYB>
A;Cross-references: EMBL:L10289
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: GTP binding; nucleotide binding; P-loop
F;40-47/Region: GTP-binding nucleotide-binding motif A (P-loop)
F;150-152/Region: GTP-binding SAK/L motif
F;269-272/Region: GTP-binding NKXD motif

Query Match     79.0%; Score 1469.5; DB 2; Length 354;
Best Local Similarity   77.4%; Pred. No. 1.2e-98;
Matches 274; Conservative  39; Mismatches  40; Indels  1; Gaps  1;

Qy    7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
Db    1 MACCLSEEAKEQKRINQEIEKQLRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 60

Qy   67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
Db   61 YSEEDRKGFEKIVYQNIFSAIQTLIAAMETLSLEYKDPSNNEHAEFLNSIDADSADIFED 120

Qy  127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 186
Db  121 GHVTAIKGCWTDPGMQECYDRRREYQLTDSAKYYLDDVERIHEPGYIPTLQDILRVVPT 180

Qy  187 TGIIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLVESDN-E 245
Db  181 TGIIEYPFDLYSIIFRMVDVGGQRSERKWIHCFENVTSIMFLVALSEYDQVLVESDNEE 240

Qy  246 NRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAA 305
Db  241 NRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMTSHLADYFPDVGDGPKCDYEAA 300

Qy  306 REFILKRMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
Db  301 REFMMDSYMDLNEDKEKMLYYHYTCATDTENIRFVFAAVKDTILQLNLKEYNLV 354

RESULT 14
JN0015
GTP-binding regulatory protein dgq alpha chain - fruit fly (Drosophila melanogaster)
N;Alternate names: dgq protein
C;Species: Drosophila melanogaster
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
C;Accession: JN0015
R;Lee, Y.J.; Dobbs, M.B.; Verardi, M.L.; Hyde, D.R.
Neuron 5, 889-898, 1990
A;Title: dgq: a drosophila gene encoding a visual system-specific G alpha molecule.
A;Reference number: JN0015; MUID:9109780I; PMID:2125225
A;Accession: JN0015
A;Molecule type: DNA
A;Residues: 1-360 <LEE>
C;Genetics:
A;Gene: dgq
A;Cross-references: FlyBase:FBgn0004435
A;Introns: 40/1; 101/3; 153/2; 196/2; 239/3; 298/1; 334/1
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Keywords: alternative splicing; GTP binding; nucleotide binding; P-loop
F;1-360/Product: GTP-binding regulatory protein dgq alpha chain I #status predicted <MAT
F;1-297/Product: GTP-binding regulatory protein dgq alpha chain II #status predi
F;40-47/Region: nucleotide-binding motif A (P-loop)
F;150-152/Region: GTP-binding SAK/L motif
F;268-271/Region: GTP-binding NKXD motif

Query Match     76.6%; Score 1425.5; DB 2; Length 360;
Best Local Similarity   75.2%; Pred. No. 1.8e-95;
Matches 270; Conservative  37; Mismatches  45; Indels  7; Gaps  1;
```

```
Qy    7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
Db    1 MACCLSEEAKEQKRINQEIEKQLRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 60

Qy   67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
Db   61 YSDEDKRGYIKLIVFQNIFMAMQSMIKAMDMLKISYGQGEHSELADLVMSIDYETVTFED 120

Qy  127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 186
Db  121 PYLNAIKTLWDDAGIQECYDRRREYQLTDSAKYYLSDLARIEQADYLPFEQDILRAVPT 180

Qy  187 TGIIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLVESDN-E 245
Db  181 TGIIEYPFDLYSIIFRMVDVGGQRSERKWIHCFENVTSIMFLVALSEYDQVLVESDNEE 240

Qy  247 RMEESKALFRTIITYPWFQNSSVILFLNKK------DLLEEKIMYSHVDYFPEYDGPQ 299
Db  241 RMEESKALFRTIITYPWFQNSSVILFLNKKIDLLEEKIMYSHVDYFPEYDGPK 300

Qy  301 QDHAAAKQDVLKKYLACNPDPEROCYSHFTTATDTENIKLVFCAVKDTIMQNALKEFNL 359
Db  301 QDHAAAKQDVLKKYLACNPDPEROCYSHFTTATDTENIKLVFCAVKDTIMQNALKEFNL 359

RESULT 15
B41534
GTP-binding protein alpha-15 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 02-Feb-2001
C;Accession: B41534
R;Wilkie, T.M.; Scherle, P.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053, 1991
A;Title: Characterization of G-protein alpha subunits in the G-q class: expression in
A;Reference number: A41534; MUID:92052208; PMID:1946421
A;Accession: B41534
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <WIL>
C;Superfamily: GTP-binding regulatory protein Gs alpha chain
C;Cross-references: GB:M80632; NID:g193570; PIDN:AAA37713.1; PID:g193571
C;Keywords: GTP binding; nucleotide binding; P-loop
F;49-56/Region: GTP-binding nucleotide-binding motif A (P-loop)
F;277-280/Region: GTP-binding NKXD motif

Query Match     57.3%; Score 1066; DB 2; Length 374;
Best Local Similarity   57.2%; Pred. No. 1.6e-69;
Matches 207; Conservative  55; Mismatches  88; Indels 12; Gaps  2;

Qy   10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSYSD 69
Db   13 CLFEEEKTAARIDQEINRILLEQKKQEREELKLLLLGPGESGKSTFLKQMRIIHGVGYSE 72

Qy   70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
Db   73 EDRRAFRLLIYQNIFVSMQAMIDAMDRLQIPFSRPDSKQHASLVMTQDPYKVSTFEKPYA 132

Qy  130 DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
Db  133 VAMQYLWRDAGIRACYERRREFHLLDSAVYYLSHLERISEDSYIPTAQDVLRSRMPTTGI 192

Qy  190 IEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLVESDNERME 249
Db  193 NEYCFSVKKTKLRIVDVGGQRSERKKWIHCFENVIALIYLASLSEYDQCLEENDQENRME 252

Qy  250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
Db  253 ESLALFSTILELPWFKSTSVILFLNKTDILEDKIHTSHLATYFPSFQGPKRDAEAAKSFI 312

Qy  310 LKMFVDL-----NPDSD----KIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
Db  313 LDMYARVYASCAEPQDGGRKGSRARRFFAHFTCATDTQSVRSVFKDVRDSVLARYLDEIN 372
```

Qy      358 LV 359
          |:
Db      373 LL 374

Search completed: June 30, 2003, 16:18:23
Job time : 20 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:   June 30, 2003, 16:10:20 ; Search time 11 seconds
(without alignments)
1353.637 Million cell updates/sec

Title:   US-09-899-295-2
Perfect score: 1860
Sequence: 1 MTLESIMACCLSEEAKEARR..........VFRAVKDTILQLNLKEYNLV 359

Scoring table:   BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:   112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters:   112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | | Description |
|---|---|---|---|---|---|---|---|
| 1 | 1830 | 98.4 | 353 | 1 | GBQ_MOUSE | P21279; | mus musculu |
| 2 | 1822 | 98.0 | 353 | 1 | GBQ_CANFA | O28294 | canis famil |
| 3 | 1818 | 97.7 | 353 | 1 | GBQ_RAT | P82471 | rattus norv |
| 4 | 1817 | 97.7 | 353 | 1 | GBQ_HUMAN | P50148 | homo sapien |
| 5 | 1766 | 94.9 | 353 | 1 | GBQ_XENLA | P38410 | xenopus lae |
| 6 | 1709 | 91.9 | 359 | 1 | GB11_MELGA | P45645 | meleagris g |
| 7 | 1691 | 90.9 | 359 | 1 | GB11_BOVIN | P38409 | bos taurus |
| 8 | 1690 | 90.9 | 359 | 1 | GB11_HUMAN | P29992 | homo sapien |
| 9 | 1685 | 90.6 | 359 | 1 | GB11_XENLA | P43444 | xenopus lae |
| 10 | 1682 | 90.4 | 359 | 1 | GB11_MOUSE | P21278 | mus musculu |
| 11 | 1674 | 90.0 | 359 | 1 | GB11_RAT | Q9Jld2 | rattus norv |
| 12 | 1538 | 82.7 | 355 | 1 | GB14_BOVIN | P38408 | bos taurus |
| 13 | 1537 | 82.6 | 354 | 1 | GB14_XENLA | O73819 | xenopus lae |
| 14 | 1535 | 82.5 | 355 | 1 | GB14_MOUSE | P30677 | mus musculu |
| 15 | 1530 | 82.3 | 353 | 1 | GBQ_PATYE | O15975 | patinopecte |
| 16 | 1529 | 82.2 | 353 | 1 | GBQ_HOMAM | P91950 | homarus ame |
| 17 | 1528 | 82.2 | 355 | 1 | GB14_HUMAN | O95837 | homo sapien |
| 18 | 1524 | 81.9 | 353 | 1 | GBQ_LYMST | P38411 | lymnaea sta |
| 19 | 1469.5 | 79.0 | 354 | 1 | GBQ_LOLFO | P38412 | loligo forb |
| 20 | 1439 | 77.4 | 353 | 1 | GBQ_DROME | P23625 | drosophila |
| 21 | 1066 | 57.3 | 374 | 1 | GB15_MOUSE | P30678 | mus musculu |
| 22 | 1052 | 56.6 | 374 | 1 | GB15_RAT | P30679 | rattus norv |
| 23 | 1044 | 56.1 | 374 | 1 | GB15_HUMAN | P30679 | homo sapien |
| 24 | 908 | 48.8 | 352 | 1 | GBA1_EMENI | Q00743 | emericella |
| 25 | 904 | 48.6 | 198 | 1 | GBA1_CANFA | P52206 | canis famil |
| 26 | 896 | 48.2 | 352 | 1 | GBA1_COCHE | O74227 | cochliobolu |
| 27 | 893.5 | 48.0 | 353 | 1 | GBA1_HUMAN | P04898 | homo sapien |
| 28 | 892.5 | 48.0 | 353 | 1 | GBA1_RAT | P10824 | rattus norv |
| 29 | 891 | 47.9 | 352 | 1 | GBA1_CRYPA | Q00580 | cryphonectr |
| 30 | 890 | 47.8 | 353 | 1 | GBA1_NEUCR | Q05425 | neurospora |
| 31 | 889 | 47.8 | 352 | 1 | GBA1_SPOSC | O74259 | sporothrix |
| 32 | 886.5 | 47.7 | 353 | 1 | GBA1_CAVPO | P38401 | cavia porce |
| 33 | 886 | 47.6 | 352 | 1 | GBA1_MAGGR | O13315 | magnaporthe |
| 34 | 884 | 47.5 | 352 | 1 | GBA1_COLTR | O42784 | colletotric |
| 35 | 883.5 | 47.5 | 353 | 1 | GBI1_CHICK | P50146 | gallus gall |
| 36 | 883.5 | 47.5 | 353 | 1 | GBI1_XENLA | P27044 | xenopus lae |
| 37 | 881.5 | 47.4 | 353 | 1 | GB02_HUMAN | P29777 | homo sapien |
| 38 | 878.5 | 47.2 | 353 | 1 | GBAK_CAVPO | P38403 | cavia porce |
| 39 | 877.5 | 47.2 | 353 | 1 | GB02_CRILO | P17806 | cricetulus |
| 40 | 877.5 | 47.2 | 353 | 1 | GBAK_RAT | P08753 | rattus norv |
| 41 | 877.5 | 47.2 | 353 | 1 | GBT2_MOUSE | P50149 | mus musculu |
| 42 | 877 | 47.2 | 354 | 1 | GBT2_CHICK | P50147 | gallus gall |
| 43 | 876.5 | 47.1 | 353 | 1 | GBAK_HUMAN | P08754 | homo sapien |
| 44 | 874 | 47.0 | 354 | 1 | GBI2_CAVPO | P38402 | cavia porce |
| 45 | 874 | 47.0 | 354 | 1 | GBI2_ORYLA | O13055 | oryzias lat |

## ALIGNMENTS

RESULT 1

GBQ_MOUSE

ID   GBQ_MOUSE     STANDARD;    PRT;    353 AA.
AC   P21279;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(q), alpha subunit.
GN   GNAQ.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=91067657; PubMed=2123549;
RA   Strathmann M., Simon M.I.;
RT   "G protein diversity: a distinct class of alpha subunits is present
RT   in vertebrates and invertebrates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:9113-9117(1990).
RN   [2]
RP   PALMITOYLATION.
RC   TISSUE=Brain;
RX   MEDLINE=94043367; PubMed=8227063;
RA   Wedegaertner P.B., Chu D.H., Wilson P.T., Levis M.J., Bourne H.R.;
RT   "Palmitoylation is required for signaling functions and membrane
RT   attachment of Gq alpha and Gs alpha.";
RL   J. Biol. Chem. 268:25001-25008(1993).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC     INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC     SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC     THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   -!-
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; M55412; AAA63306.1; ALT_INIT.
DR   PIR; A38414; RGMSQ.
DR   HSSP; P10824; 1BOF.
DR   MGD; MGI:95776; Gnaq.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW   Palmitate; Lipoprotein.
FT   LIPID     3    3    PALMITATE.

FT   LIPID                             PALMITATE.
FT   NP_BIND        40     47          GTP (BY SIMILARITY).
FT   NP_BIND       199    203          GTP (BY SIMILARITY).
FT   NP_BIND       268    271          GTP (BY SIMILARITY).
FT   MOD_RES       177    177          ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT                                     (BY SIMILARITY).
SQ   SEQUENCE   353 AA;  41478 MW;  C41B2AC11C674C5F CRC64;

Query Match     98.4%;  Score 1830;  DB 1;  Length 353;
Best Local Similarity   99.7%;  Conservative 1;  Pred. No. 2.5e-130;
Matches 352;                     Mismatches 0;  Indels 0;  Gaps 0;

```
Qy     7  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMRIIHGSG  66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMRIIHGSG  60

Qy    67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  120

Qy   127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  180

Qy   187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy   247  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  300

Qy   307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  353
```

RESULT 2
GBQ_CANFA

ID   GBQ_CANFA     STANDARD;      PRT;   353 AA.
AC   Q28294;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(q), alpha subunit.
GN   GNAQ.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96433124; PubMed=8836152;
RA   Johnson G.J., Leis L.A., Dunlop P.C.;
RT   "Specificity of G alpha q and G alpha 11 gene expression in platelets
RT   and erythrocytes. Expressions of cellular differentiation and species
RT   differences.";
RL   Biochem. J. 318:1023-1031(1996).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC   -!- SIMILARITY: THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).

DR   EMBL; L76257; AAB05548.1; ALT_INIT.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   PRINTS; PR00503; G-alpha; 1.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW   Palmitate; Lipoprotein.
FT   LIPID           3      3        PALMITATE (BY SIMILARITY).
FT   LIPID           4      4        PALMITATE (BY SIMILARITY).
FT   NP_BIND        40     47        GTP (BY SIMILARITY).
FT   NP_BIND       199    203        GTP (BY SIMILARITY).
FT   NP_BIND       268    271        GTP (BY SIMILARITY).
FT   MOD_RES       177    177        ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT                                   (BY SIMILARITY).
SQ   SEQUENCE   353 AA;  41467 MW;  E7737E07B1F4904C CRC64;

Query Match     98.0%;  Score 1822;  DB 1;  Length 353;
Best Local Similarity   99.4%;  Conservative 1;  Pred. No. 9.9e-130;
Matches 351;                    Mismatches 1;  Indels 0;  Gaps 0;

```
Qy     7  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTGESGKSTFIKQMRIIHGSG  66
          |||||||||||||||||||||:|::||||||||||||||||||||||||||||||||||
Db     1  MACCLSEEAKEARRINDEIERLRRDKRDARRELKLLLGTGESGKSTFIKQMRIIHGSG   60

Qy    67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  120

Qy   127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  180

Qy   187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy   247  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  300

Qy   307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  353
```

RESULT 3
GBQ_RAT

ID   GBQ_RAT      STANDARD;      PRT;   353 AA.
AC   P82471;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(q), alpha subunit.
GN   GNAQ.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strotmann R.;
RT   "Rat G alpha q subunit.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 74-229 FROM N.A.
RC   TISSUE=Brain cortex;
RX   MEDLINE=93081611; PubMed=1333286;
RA   Lesch K.-P., Manji H.K.;
RT   "Signal-transducing G proteins and antidepressant drugs; evidence for
RT   modulation of alpha subunit gene expression in rat brain.";

RESULT 4
ID   GBO_HUMAN     STANDARD;     PRT;     353 AA.

```
RL   Biol. Psychiatry 32:549-579(1992).
RN   [3]
RP   SEQUENCE OF 238-331 FROM N.A.
RA   Thomas C.P.;
RT   "GTP-binding protein expression in glomerular mesangial cells.";
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   ---------------------------------------------------------------------------
```
```
DR   EMBL; AF234260; AAF59930.1; ALT_INIT.
DR   EMBL; L37294; AAB02848.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW   Palmitate; Lipoprotein.
FT   LIPID         3      3       PALMITATE (BY SIMILARITY).
FT   LIPID         4      4       PALMITATE (BY SIMILARITY).
FT   NP_BIND      40     47       GTP (BY SIMILARITY).
FT   NP_BIND     199    203       GTP (BY SIMILARITY).
FT   NP_BIND     268    271       GTP (BY SIMILARITY).
FT   MOD_RES     177    177       ADP-RIBOSYL(1) (BY SIMILARITY).
FT                                ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT   CONFLICT     85     85       V -> I (IN REF. 2).
SQ   SEQUENCE   353 AA;   41469 MW;   BB4C21FDDD47534 CRC64;
```

```
Query Match        349;
Best Local Similarity   97.78%;
Matches  349;   Conservative  3;   Mismatches 1;   Indels 0;   Gaps 0;
                   Score 1818;   DB 1;   Length 353;
                   98.9%;   Pred. No. 2e-129;

Qy      7  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG   66
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1  MACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG   60

Qy     67  YSDEDKRGTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
           ||||||||||||||||||||||::||||||||||||||||||||||||||||||||||
Db     61  YSDEDKRGTKLVYQNIFTAMQAVRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN   120

Qy    127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYILNDLDRVADPAYLPTQQDVLRVRVPT  186
           |||||||||||||||||||||:||||||||||||||||||||||:||||||||||||||
Db    121  PYVDAIKSLWNDPGIQECYDRREYQLSDSTKYILNDLDRVADPSYLPTQQDVLRVRVPT   180

Qy    187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy    247  RMEESKALFRTIIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  RMEESKALFRTIIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  300

Qy    307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  353
```

```
AC   P50148; Q13462; Q92471; O15108;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(q), alpha subunit.
GN   GNAQ OR GAQ.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96423032; PubMed=8825633;
RA   Dong Q., Shenker A., Way J., Haddad B.R., Lin K., Hughes M.R.,
RA   McBride W.O., Spiegel A.M., Battey J.;
RT   "Molecular cloning of human G alpha q cDNA and chromosomal
RT   localization of the G alpha q gene (GNAQ) and a processed
RT   pseudogene.";
RL   Genomics 30:470-475(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Prostate;
RX   MEDLINE=96256639; PubMed=8664309;
RA   Chen B., Leverette R.D., Schwinn D.A., Kwatra M.M.;
RT   "Human G(alpha q): cDNA and tissue distribution.";
RL   Biochim. Biophys. Acta 1281:125-128(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96433124; PubMed=8836152;
RA   Johnson G.J., Leis L.A., Dunlop P.C.;
RT   "Specificity of G alpha q and G alpha 11 gene expression in platelets
RT   and erythrocytes. Expressions of cellular differentiation and species
RT   differences.";
RL   Biochem. J. 318:1023-1031(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Bai X.H., Acharya R., Rivera C., Murtagh J.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 74-229 FROM N.A.
RC   TISSUE=Brain cortex;
RX   MEDLINE=93081611; PubMed=1333286;
RA   Lesch K.-P., Manji H.K.;
RT   "Signal-transducing G proteins and antidepressant drugs: evidence for
RT   modulation of alpha subunit gene expression in rat brain.";
RL   Biol. Psychiatry 32:549-579(1992).
RN   [6]
RP   SEQUENCE OF 238-331 FROM N.A.
RC   TISSUE=Hematopoietic;
RX   MEDLINE=96077138; PubMed=7492305;
RA   Thomas C.P., Dunn M.J., Mattera R.;
RT   "Ca2+ signalling in K562 human erythroleukaemia cells: effect of
RT   dimethyl sulphoxide and role of G-proteins in thrombin- and
RT   thromboxane A2-activated pathways.";
RL   Biochem. J. 312:151-158(1995).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, PROSTATE,
CC       TESTIS AND COLON.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   ---------------------------------------------------------------------------
```
```
DR   EMBL; U40038; AAC50363.1; ALT_INIT.
```

## Left column

```
DR   EMBL; U43083; AAB06875.1; ALT_INIT.
DR   EMBL; L76256; AAB39498.1; ALT_INIT.
DR   EMBL; AF011496; AAB64301.1; ALT_INIT.
DR   EMBL; L40629; AAA99950.1; -.
DR   Genew; HGNC:4390; GNAQ.
DR   HSSP; P10824; 1BOF.
DR   MIM; 600998;
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW   Palmitate; Lipoprotein.
FT   LIPID         3      3       PALMITATE (BY SIMILARITY).
FT   LIPID         4      4       PALMITATE (BY SIMILARITY).
FT   NP_BIND      40     47       GTP (BY SIMILARITY).
FT   NP_BIND     199    203       GTP (BY SIMILARITY).
FT   NP_BIND     268    271       GTP (BY SIMILARITY).
FT   MOD_RES     177    177       ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT                                (BY SIMILARITY).
FT   CONFLICT     22     23       QL -> HV (IN REF. 1).
FT   CONFLICT     86     86       R -> T (IN REF. 4).
FT   CONFLICT     97     97       Y -> C (IN REF. 4).
FT   CONFLICT    165    165       A -> S (IN REF. 5).
FT   CONFLICT    318    318       I -> V (IN REF. 3).
FT   CONFLICT    331    331       I -> V (IN REF. 3).
FT   CONFLICT    352    352       A -> L (IN REF. 1 AND 4).
SQ   SEQUENCE   353 AA;  41425 MW;  E7798E07B1F49D4C CRC64;

Query Match             97.7%;
Best Local Similarity   99.2%;
Matches  350;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;
                        Score 1817;  Pred. No. 2.3e-129;  DB 1;  Length 353;

Qy        7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        1 MACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 60

Qy       67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 120

Qy      127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 186
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 180

Qy      187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 246
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 240

Qy      247 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 306
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 300

Qy      307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNKEYNLV 359
            ||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      301 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNKEYNAV 353


RESULT 5
GBQ_XENLA
ID   GBQ_XENLA    STANDARD;    PRT;    353 AA.
AC   P38410;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(q), alpha subunit.
GN   GNAQ.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
```

## Right column

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ladner K.J., Smith L.D.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocyte;
RP   MEDLINE=94298961; PubMed=8026589;
RA   Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RA   Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RT   "Neuromedin B receptor, expressed in Xenopus laevis oocytes,
RT   selectively couples to G alpha q and not G alpha 11.";
RL   FEBS Lett. 348:89-92(1994).
RN   [3]
RA   Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RL   FEBS Lett. 349:318-318(1994).
-!- ERRATUM.
-!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
    INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
    SIGNALING SYSTEMS.
-!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
    THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
-!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY.
-!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
DR   EMBL; L05540; AAA49730.1; ALT_INIT.
DR   EMBL; U10502; AAA52189.1; ALT_INIT.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW   Palmitate; Lipoprotein.
FT   LIPID         4      4       PALMITATE (BY SIMILARITY).
FT   LIPID         3      3       PALMITATE (BY SIMILARITY).
FT   NP_BIND      40     47       GTP (BY SIMILARITY).
FT   NP_BIND     199    203       GTP (BY SIMILARITY).
FT   NP_BIND     268    271       GTP (BY SIMILARITY).
FT   MOD_RES     177    177       ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT                                (BY SIMILARITY).
FT   CONFLICT     10     10       E -> K (IN REF. 2).
FT   CONFLICT     79     79       S -> T (IN REF. 2).
FT   CONFLICT    158    158       L -> V (IN REF. 2).
FT   CONFLICT    164    164       H -> Q (IN REF. 2).
SQ   SEQUENCE   353 AA;  41541 MW;  23101424121184553 CRC64;

Query Match             94.9%;
Best Local Similarity   96.08%;
Matches  339;  Conservative   7;  Mismatches   7;  Indels   0;  Gaps   0;
                        Score 1766;  Pred. No. 1.6e-125;  DB 1;  Length 353;

Qy        7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        1 MACCLSEEAKEARRINDEIEROLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 60

Qy       67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
            ||||||||||||||||||:|||||||||||:|||||||||||:||:||||||||||||||
Db       61 YSDEDKRGFTKLVYQNIFSAMQAMIRAMETLKIPYKYEHNKGHALLVREVDVEKVSAFEN 120

Qy      127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT 186
            |||||| :||||||||||||||||||||||||||||||||:|::| ||||||||||||||
Db      121 PYVDAIKLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRIATHGYLPTQQDVLRVRVPT 180

Qy      187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 246
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 240
```

```
Qy     247 RMEESKALFRTIIIYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 RMEESKALFRTIIIYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 300

Qy     307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353
```

RESULT 6
GB11_MELGA

```
ID   GB11_MELGA     STANDARD;     PRT;   359 AA.
AC   P45645;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-11 subunit.
GN   GNA11.
OS   Meleagris gallopavo (Common turkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX   NCBI_TaxID=9103;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=93207527; PubMed=8457205;
RA   Maurice D.H., Waldo G.L., Morris A.J., Nicholas R.A., Harden T.K.;
RT   "Identification of G alpha 11 as the phospholipase C-activating G-
RT   protein of turkey erythrocytes.";
RL   Biochem. J. 290:765-770(1993).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X73072; CAA51530.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND      46     53       GTP (BY SIMILARITY).
FT   NP_BIND     205    209       GTP (BY SIMILARITY).
FT   NP_BIND     274    277       GTP (BY SIMILARITY).
FT   MOD_RES     183    183       ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT                               (BY SIMILARITY).
SQ   SEQUENCE   359 AA;  42058 MW;  B9E16427169BE1AE CRC64;
```

```
Query Match             91.9%;  Score 1709;  DB 1;  Length 359;
Best Local Similarity   90.3%;  Pred. No. 3.1e-121;
Matches  324;  Conservative  22;  Mismatches  13;  Indels   0;  Gaps   0;

Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLGTEGSGKSTFIKQMR 60
          |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db      1 MTLESMMACCLSDEVKESKRINAEIEKQLRRDKRDARRELKLLLGTGESGKSTFIKQMR 60

Qy     61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||||||||||||||||||||||||||||:|||||| |||::|||| ||||||||
Db     61 IIHGSGYSEEDKGFTKLVYQNIFTAMQSMIRAMETLKILYKYEQNKANAVLIREVDVEK 120
```

---

```
Qy     121 VSAFENPYVDAIKSLWNDPGTQEECYDRRREYQLSDSTKYYLNDLLDRVADPAYLPTQQDVL 180
           | | ||| :| :|||:|| ||::||||||||||||:|||||||||||||:||::|:| |||||
Db     121 VMTFEQPYVSAIKTLWNDPGIQEECYDRRREYQLSDSAKYYLSDVDRIATPGYLPTQQDVL 180

Qy     181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
           |||||||||||||||| :::||||||||||||||||||||||||||||||||||||||||
Db     181 RVRVPTTGIIEYPFDLENIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy     241 ESDNENRMEESKALFRTIIIYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
           |||||||||||||||||||||||||||||||||||:|||||:|||||||||||||||||||
Db     241 ESDNENRMEESKALFRTIIIYPWFQNSSVILFLNKKDLLEDKILYSHLVDYFPEPDGPQR 300

Qy     301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 7
GB11_BOVIN

```
ID   GB11_BOVIN     STANDARD;     PRT;   359 AA.
AC   P38409;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-11 subunit (GL2).
GN   GNA11.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91286303; PubMed=1905731;
RA   Nakamura F., Ogata K., Shiozaki K., Kameyama K., Ohara K., Haga T.,
RA   Nukada T.;
RT   "Identification of two novel GTP-binding protein alpha-subunits that
RT   lack apparent ADP-ribosylation sites for pertussis toxin.";
RL   J. Biol. Chem. 266:12676-12681(1991).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D90336; BAA14350.1; ALT_INIT.
DR   PIR; B40891; B40891.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND      46     53       GTP (BY SIMILARITY).
FT   NP_BIND     205    209       GTP (BY SIMILARITY).
FT   NP_BIND     274    277       GTP (BY SIMILARITY).
FT   MOD_RES     183    183       ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT                               (BY SIMILARITY).
SQ   SEQUENCE   359 AA;  42070 MW;  B0BAC4FBF5AAE8D5 CRC64;
```

```
Query Match             90.9%;  Score 1691;  DB 1;  Length 359;
Best Local Similarity   89.4%;  Pred. No. 6.9e-120;
Matches  321;  Conservative  24;  Mismatches  14;  Indels   0;  Gaps   0;
```

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60
        ||||:|:|| |||:||:||:::|||| |::||||:|::||||||||||||||||||||||
Db    1 MTLESMMACCLSDEVKESKRINAEIEKQLRPDKRDARRELKLILLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
        ||||:|:|||:||||||||||||||||||||||:|||||||||:|||||:||||||||||
Db   61 IIHGAGYSEEDKRGFTKLVYQNIFTAMQAMIRAMETLKILYKYEQNKANALLIREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
        |:||::||:||:|||||||||||||||||||||:|||||||||||:||||||||||||
Db  121 VTTFEHRYVSAIKTLWNDPGIQECYDRREYQLSDSAKYYLTDVDRIATSGYLPTQQDVL 180

Qy  181 RVRVPTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
        |||||||||||||||:|:||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
        ||||||||||||||||||||||||||||||||||||||:|||||||:|||||:|||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEDKILHSHLVDYFPEFDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCADTENIRFVFAAVKDTILQLNLKEYNLV 359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCADTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 8

GB11_HUMAN

```
ID   GB11_HUMAN    STANDARD;    PRT;    359 AA.
AC   P29992; Q14350; O15109;
DT   01-APR-1993 (Rel. 25, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(Y), alpha subunit (Alpha-11).
GN   GNA11 OR GA11.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91219481; PubMed=1902575;
RA   Jiang M., Pandey S., Tran V.T., Fong H.K.W.;
RT   "Guanine nucleotide-binding regulatory proteins in retinal pigment
RT   epithelial cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:3907-3911(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Bai X.H., Acharya R., Bai Y.H., Murtagh J.J.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA   Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA   Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA   Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA   Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA   Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA   Kobayashi A., Olsen A.S., Carrano A.V.;
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 244-337 FROM N.A.
RC   TISSUE=Hematopoietic;
RX   MEDLINE=96077138; PubMed=7492305;
RA   Thomas C.P., Dunn M.J., Mattera R.;
RT   "Ca2+-signaling in K562 human erythroleukaemia cells: effect of
RT   dimethyl sulphoxide and role of G-proteins in thrombin- and
RT   thromboxane A2-activated pathways.";
RL   Biochem. J. 312:151-158(1995).
```

```
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   -------------------------------------------------------------------
```

```
DR   EMBL; M69013; AAA58624.1; -.
DR   EMBL; AF011497; AAB64303.1; -.
DR   EMBL; AC005262; AAC25615.1; -.
DR   EMBL; L40630; AAA99949.1; -.
DR   PIR; A39394; RGHUGY.
DR   HSSP; P10824; 1BOF.
DR   Genew; HGNC:4379; GNA11.
DR   MIM; 139313; -.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND      46     53       GTP (BY SIMILARITY).
FT   NP_BIND     205    209       GTP (BY SIMILARITY).
FT   NP_BIND     274    277       GTP (BY SIMILARITY).
FT   MOD_RES     183    183       ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT   CONFLICT      6      6       M -> I (IN REF. 2).
FT   CONFLICT    266    266       M -> I (IN REF. 4).
FT   CONFLICT    285    285       N -> H (IN REF. 4).
FT   CONFLICT    301    302       Y -> H (IN REF. 4).
FT   CONFLICT    310    310       DA -> EP (IN REF. 1).
FT   CONFLICT    310    310       L -> P (IN REF. 2).
SQ   SEQUENCE   359 AA;   42123 MW;   DD37176589E56046 CRC64;
```

```
Query Match             90.9%;   Score 1690;  DB 1;  Length 359;
Best Local Similarity   89.7%;   Pred. No. 8.2e-120;
Matches  322;  Conservative  22;  Mismatches  15;  Indels  0;  Gaps  0;
```

```
Qy    1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMR 60
        ||||:|:|| |||:||:||:::|||| |::||||:|::||||||||||||||||||||||
Db    1 MTLESMMACCLSDEVKESKRINAEIEKQLRPDKRDARRELKLILLGTGESGKSTFIKQMR 60

Qy   61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
        ||||:|:|||:||||||||||||||||||||||:|||||||||:|||||:||||||||||
Db   61 IIHGAGYSEEDKRGFTKLVYQNIFTAMQAMIRAMETLKILYKYEQNKANALLIREVDVEK 120

Qy  121 VSAFENPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
        |:||::||:||:|||||||||||||||||||||:|||||||||||:||||||||||||
Db  121 VTTFEHRYVSAIKTLWNDPGIQECYDRREYQLSDSAKYYLTDVDRIATSGYLPTQQDVL 180

Qy  181 RVRVPTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
        |||||||||||||||:|:||||||||||||||||||||||||||||||||||||||||
Db  181 RVRVPTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
        ||||||||||||||||||||||||||||||||||||||:|||||||:|||||:|||||
Db  241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEDKILHSHLVDYFPEFDGPQR 300

Qy  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCADTENIRFVFAAVKDTILQLNLKEYNLV 359
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCADTENIRFVFAAVKDTILQLNLKEYNLV 359
```

RESULT 9

GB11_XENLA

```
ID   GB11_XENLA    STANDARD;    PRT;    359 AA.
AC   P43444;
DT   01-NOV-1995 (Rel. 32, Created)
```

```
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-11 subunit.
GN   GNA11.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocyte;
RX   MEDLINE=94298961; PubMed=8025589;
RA   Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RL   "Neuromedin B receptor, expressed in Xenopus laevis oocytes,
RL   selectively couples to G alpha q and not G alpha 11.";
RL   FEBS Lett. 348:89-92(1994).
RN   [2]
RP   ERRATUM.
RA   Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RL   FEBS Lett. 349:318-318(1994).
-!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
    INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
    SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
-!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
    THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
-!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
```

```
DR   EMBL; U10494; AAA52188.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   PRINTS; PR00503; G-alpha; 1.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND      46     53       GTP (BY SIMILARITY).
FT   NP_BIND     205    209       GTP (BY SIMILARITY).
FT   NP_BIND     274    277       GTP (BY SIMILARITY).
FT   MOD_RES     183    183       ADP-RIBOSYL[1] (BY ACTION OF CTX) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   359 AA;  42088 MW;  DA6B376993FDDB70 CRC64;

Query Match             90.68%;  Score 1685;  DB 1;  Length 359;
Best Local Similarity   88.9%;   Pred. No. 1.9e-119;
Matches  319;  Conservative  25;  Mismatches  15;  Indels  0;  Gaps  0;

Qy    1  MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR   60
         ||||:|  |||||||||| ||||||||||||||||| :: |||||||||||||||||||||
Db    1  MTLDSTMACCLSEEVKESKRINAEIEKQLRRDKKDSRRELKILLLLGTGESGKSTFIKQMR   60

Qy   61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
         ||||||:  |||||  ||||:||||||| ||:| |  |  |: ||||||||||||||||||
Db   61  IIHGSGYSEEDKKGFTKLVFQNIFTAMQSMIRAMETLKILYKYEQNKANAQVVREVDVEK  120

Qy  121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRYADPAYLPTQODVL  180
         | ||| ||||  ||||||||| |||||||||||||:|||||:: |||| || |||| |||
Db  121  VCTFEQPYVNAIKNLWSDPGIQECYDRRREYQLSDSTKYYLIDVDRISKPGYLPTQQDVL  180

Qy  181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
         |||||||||||:|||| |||| ||||||||:||||||||||||||||:||||||||||||
Db  181  RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240

Qy  241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEDKIMYSHLVDYFPEFDGPQR  300

Qy  301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRRFVFAAVKDTILQLNLKEYNLV  359
         || ||||||||||||||||||||||||||||||||| ||||:|||||||||||||||||
Db  301  DAATAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQHNLKEYNLV  359


RESULT 10
GB11_MOUSE
ID   GB11_MOUSE    STANDARD;    PRT;    359 AA.
AC   P21278; Q61939;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-11 subunit.
GN   GNA11 OR GNA-11.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91067657; PubMed=2123549;
RA   Strathmann M., Simon M.I.;
RL   "G protein diversity: a distinct class of alpha subunits is present
RL   in vertebrates and invertebrates.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:9113-9117(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RX   MEDLINE=96435432; PubMed=8838318;
RA   Davignon I., Barnard M., Gavrilova O., Sweet K., Wilkie T.M.;
RL   "Gene structure of murine Gna11 and Gna15: tandemly duplicated Gq
RL   class G protein alpha subunit genes.";
RL   Genomics 31:359-366(1996).
RN   [3]
RP   SEQUENCE OF 211-271 FROM N.A.
RX   MEDLINE=90017488; PubMed=2508088;
RA   Strathmann M., Wilkie T.M., Simon M.I.;
RL   "Diversity of the G-protein family: sequences from five additional
RL   alpha subunits in the mouse.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:7407-7409(1989).
-!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
    INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
    SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
-!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
    THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
-!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
```

```
DR   EMBL; M55411; AAA63305.1; -.
DR   EMBL; U37411; AAB36839.1; -.
DR   EMBL; U37411; AAB36839.1; JOINED.
DR   EMBL; U37412; AAB36839.1; JOINED.
DR   EMBL; M57617; AAA63301.1; -.
DR   PIR; B38414; RGMS11.
DR   PIR; B33833; B33833.
DR   HSSP; P10824; 1BOF.
DR   MGD; MGI:95766; Gna11.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   PRINTS; PR00503; G-alpha; 1.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND      46     53       GTP (BY SIMILARITY).
```

FT   CONFLICT    77     77
FT   NP_BIND    205    209       GTP (BY SIMILARITY).
FT   NP_BIND    274    277       GTP (BY SIMILARITY).
FT   NP_BIND    183    183       ADP-RIBOSYL(1) (BY ACTION OF CTX)
FT   MOD_RES                     (BY SIMILARITY).
FT   SQ  SEQUENCE   359 AA;  42024 MW;  A33D2D6C6C62F8D5 CRC64;
FT   CONFLICT    77     77       K -> L (IN REF. 2).

Query Match             90.4%;  Score 1682;  DB 1;  Length 359;
Best Local Similarity   88.6%;  Pred. No. 3.3e-119;
Matches 318;  Conservative 28;  Mismatches 13;  Indels 0;  Gaps 0;

Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||||||||||::|||||||||||||||||||||||||||
Db      1 MTLESMMACCLSDEVKESKRINAEIEKQLRDKDRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy     61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          |||||||:|||||||||||||||||||||||||||:|||| |||||||||| |||||||||
Db     61 IIHGAGYSEEDKRGFTKLVYQNIFTAMQAVVRAMDTLKIRYKYEQNKANALLIREVDVEK 120

Qy    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |:|:||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db    121 VTTFEHQYVNAIKTLWSDPGVQECYDRRREFQLSDSAKYLTDVDRIATVGYLPTQQDVL 180

Qy    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db    181 RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||:|::||||||||||||||||||||
Db    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEDKILHSHLVDYFPEFDGPQR 300

Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

---

RESULT 11
ID   GB11_RAT        STANDARD;      PRT;   359 AA.
AC   Q9JID2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-11 subunit.
GN   GNA11.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strotmann R.;
RT   "Rattus norvegicus guanine nucleotide binding protein alpha 11 subunit
RT   (G11).";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF239674; AAF81690.1; -.
DR   HSSP; P10824; 1BOF.

DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND     46     53       GTP (BY SIMILARITY).
FT   NP_BIND    205    209       GTP (BY SIMILARITY).
FT   NP_BIND    274    277       GTP (BY SIMILARITY).
FT   NP_BIND    183    183       ADP-RIBOSYL(1) (BY ACTION OF CTX)
FT   MOD_RES                     (BY SIMILARITY)
SQ   SEQUENCE   359 AA;  42026 MW;  B4CD0578F9C7092A CRC64;

---

RESULT 12
ID   GB14_BOVIN      STANDARD;      PRT;   355 AA.
AC   P38408;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-14 subunit (G11).
GN   GNA14.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91286303; PubMed=1905731;
RA   Nakamura F., Ogata K., Shiozaki K., Kameyama K., Ohara K., Haga T.,
RA   Nukada T.;
RT   "Identification of two novel GTP-binding protein alpha-subunits that
RT   lack apparent ADP-ribosylation sites for pertussis toxin.";
RL   J. Biol. Chem. 266:12676-12681(1991).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -

Query Match             90.0%;  Score 1674;  DB 1;  Length 359;
Best Local Similarity   88.0%;  Pred. No. 1.3e-118;
Matches 316;  Conservative 30;  Mismatches 13;  Indels 0;  Gaps 0;

Qy      1 MTLESIMACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR 60
          ||||||||||||||||||||||||::|||||||||||||||||||||||||||||||||||
Db      1 MTLESMIACCLSDEVKESKRINAEIEKQLRDKDRDARRELKLLLLGTGESGKSTFIKQMR 60

Qy     61 IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK 120
          ||||||||:||||||||||||||||||||||||||||:|||| |||||||||| ||||||||
Db     61 IIHGAGYSEEDKRGFTKLVYQNIFTAMQAVVRAMDTLKIRYKYEQNKANALLIREVDVEK 120

Qy    121 VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL 180
          |:|:||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db    121 VTTFEHQYVNAIKTLWSDPGVQECYDRRREFQLSDSAKYLTDVDRIATVGYLPTQQDVL 180

Qy    181 RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240
          |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db    181 RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV 240

Qy    241 ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR 300
          |||||||||||||||||||||||||||||||||||||:|::||||||||||||||||||||
Db    241 ESDNENRMEESKALFRTIITYPWFQHSSVILFLNKKDLLEDKILHSHLVDYFPEFDGPQR 300

Qy    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359

**Left column**

```
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; D90335; BAA14349.1; -.
DR   PIR; A40891; A40891.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND    42     49   GTP (BY SIMILARITY).
FT   NP_BIND   201    205   GTP (BY SIMILARITY).
FT   NP_BIND   270    273   GTP (BY SIMILARITY).
FT   NP_BIND   179    179   GTP (BY SIMILARITY).
FT   MOD_RES   179    179   ADP-RIBOSYL(1) (BY ACTION OF CTX)
FT                          (BY SIMILARITY).
SQ   SEQUENCE   355 AA;  41498 MW;  60888C26C9B7243B CRC64;

Query Match             82.7%;  Score 1538;  DB 1;  Length 355;
Best Local Similarity   82.3%;  Pred. No. 2.1e-108;
Matches  289;  Conservative   33;  Mismatches   29;  Indels    0;  Gaps    0;

Qy    9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 68
          || | ||||| || |||| | ||||||||||||||||||||||||||||||||||||
Db    5 CCLSAEEKESQRISAEIERQLRDKRDKDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 64

Qy   69 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPIKYKEHNKAHAQLVREVDVEKVSAFENPY 128
          ||| |||||| |||||||||||||||  ||| ||  ||||||||||||||  ||||||||
Db   65 DEDRKGFTKLVIQNIFTAMQAMIRAMDFLKIQIYVCEQNKENAQLIREVEVDKVSTLSRDQ 124

Qy  129 VDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTG 188
          | ||  | |||||||| ||||||||| |||||  ||||||||||||||| ||||| |||
Db  125 VEAIKQLWQDPGIQECYDRRREYQLSDSAKYYLTDIDRIAMPAFVPTQQDVLRVRVPTTG 184

Qy  189 IIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
          ||||||| | |||||||||||||||||||||||||| | ||||||||||| |||||||
Db  185 IIEYPFDLENIIFRMVDVGGQQRSERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244

Qy  249 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREF 308
          ||||||  |||||||  |||||||||||| ||||||||||  || ||  |||||||| |
Db  245 EESKALFKTIITYPWFLNSSVILFLNKKDLEEKIMYSHLISYFPEYTGPKQDVKAARDF 304

Qy  309 ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||| ||||||  ||  | |||||||||||||||||||| || | | ||||
Db  305 ILKLYQDQNPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
```

```
RESULT 13
ID   GB14_XENLA     STANDARD;      PRT;   354 AA.
AC   O73819;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-14 subunit.
GN   GNA14.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98344029; PubMed=9677362;
RA   Shapira H., Amit I., Revach M., Oron Y., Battey J.F.;
RT   "Galpha14 and Galphaq mediate the response to trypsin in Xenopus
RT   oocytes.";
RL   J. Biol. Chem. 273:19431-19436(1998).
```
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE

**Right column**

```
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC       MEDIATES RESPONSES TO TRYPSIN.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; AF059182; AAC41382.1; -.
DR   SMART; SM00275; G-alpha; 1.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   Pfam; PF00503; G-alpha; 1.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   HSSP; P10824; 1BOF.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND    41     48   GTP (BY SIMILARITY).
FT   NP_BIND   200    204   GTP (BY SIMILARITY).
FT   NP_BIND   269    272   GTP (BY SIMILARITY).
FT   NP_BIND   178    178   GTP (BY SIMILARITY).
FT   MOD_RES   178    178   ADP-RIBOSYL(1) (BY ACTION OF CTX)
FT                          (BY SIMILARITY).
SQ   SEQUENCE   354 AA;  41595 MW;  C71050026B037600E CRC64;

Query Match             82.6%;  Score 1537;  DB 1;  Length 354;
Best Local Similarity   80.9%;  Pred. No. 2.5e-108;
Matches  284;  Conservative   43;  Mismatches   24;  Indels    0;  Gaps    0;

Qy    9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS 68
          ||| | ||   || |||| | ||||||||||||||||||||||||||||||||||||
Db    4 CCLSAEEKESQRINAEIEKQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYT 63

Qy   69 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPIKYKEHNKAHAQLVREVDVEKVSAFENPY 128
          ||| |||||||||||||  | ||||||| || |   ||| ||| |||||| |||| |
Db   64 DEDRKGFTKLVYQNIFTMSQSMIRAMDTLRIQYTSEQNMENALVIREVDVKVSSLERKH 123

Qy  129 VDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTG 188
          | ||  || |  |||||||||||||||||||||| |||   ||  |||||||||||||||
Db  124 VEAIKKLWEDEGIQECYDRRREYQLSDSTKYYLSDIDRISNPGFIPTQQDVLRVRVPTTG 183

Qy  189 IIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
          ||||||| | |||||||||||||||||||||||||| | |||||||||||| |||||||
Db  184 IIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIIFLVALSEYDQVLAECDNENRM 243

Qy  249 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREF 308
          ||||||  ||||||||||||||||||||| |||||||||| |||| |  ||| |||| |
Db  244 EESKALFKTIITYPWFQNSSVILFLNKKDLQEKIMYSHLIDYFPEFTGPKQDSQAARDF 303

Qy  309 ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
          ||| | | ||  ||  | |||||||||||||||||||| || | | ||||
Db  304 ILKLYQDQNPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 354
```

```
RESULT 14
ID   GB14_MOUSE     STANDARD;      PRT;   355 AA.
AC   P30677;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein, alpha-14 subunit.
GN   GNA14 OR GNA-14.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090.
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92052208; PubMed=1946421;
RA   Wilkie T.M., Scherle P.A., Strathmann M.P., Slepak V.Z., Simon M.I.;
RT   "Characterization of G-protein alpha subunits in the Gq class:
RT   expression in murine tissues and in stromal and hematopoietic cell
RT   lines.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:10049-10053(1991).
RN   [2]
RP   SEQUENCE OF 217-267 FROM N.A.
RX   MEDLINE=90017488; PubMed=2508088;
RA   Strathmann M., Wilkie T.M., Simon M.I.;
RT   "Diversity of the G-protein family: sequences from five additional
RT   alpha subunits in the mouse.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:7407-7409(1989).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   -----------------------------------------------------------------------
```
```
CC   -----------------------------------------------------------------------
DR   EMBL; M80631; AAA83222.1; -.
DR   EMBL; M57616; AAA63304.1; -.
DR   PIR; A41534; E33833.
DR   PIR; E33833; E33833.
DR   HSSP; P10824; 1BOF.
DR   MGD; MGI:95769; Gna14.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT   NP_BIND       42     49       GTP (BY SIMILARITY).
FT   NP_BIND      201    205       GTP (BY SIMILARITY).
FT   NP_BIND      270    273       GTP (BY SIMILARITY).
FT   MOD_RES      179    179       ADP-RIBOSYL(1) (BY ACTION OF CTX)
SQ   SEQUENCE   355 AA;  41522 MW;  A4810D72169878E9 CRC64;

Query Match       82.5%;  Score 1535;  DB 1;  Length 355;
Best Local Similarity  82.1%;  Pred. No. 3.5e-108;
Matches 288;  Conservative 35;  Mismatches 28;  Indels 0;  Gaps 0;

Qy      9  CCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS  68
           ||| :|||:||||||||||||::|| ||||||||||||||||||||||||||||||||||
Db      5  CCLSAEEKESQRISAEIERHVRRDKDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYS  64

Qy     69  DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPY  128
           |||||||||||||||||||||||||||||:| :||  ::||:||:||:||:|:|| :| |
Db     65  DEDKRGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ  124

Qy    129  VDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTG  188
           | |||.||.||||||||||||||||||||.:|| |:||||:|| | ::::|||| | |||
Db    125  VAAIKQLWLDPGIQECYDRRREYQLSDSAXYLTDIERIAMPSFVPTQQDVLRVPTTG  184

Qy    189  IIEYPFDLQSVIFRMVDVGGQRSERKKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM  248
           ||||||||:.:|||||||||||||||||||||||.||||.||||||||||| | ||||||
Db    185  IIEYPFDLENIIFRMVDVGGQRSERKKWIHCFESVTSIIFLVALSEYDQVLAECNENRM  244

Qy    249  EESRALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREF  308
           |||||||||||||||| |||||||||||||||||||||||:.:||||:|| ::: ||| |
Db    245  EESRALFRTIITYPWFLNSSVILFLNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAARDF  304

Qy    309  ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
           |||: : |:||.||:|||||||||||||||||||||||||||||||:| ||
Db    305  ILKLYQBDQNPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV  355
```

RESULT 15

```
Db   305  ILKLYQBDQNPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV  355

ID   GBQ_PATYE      STANDARD;      PRT;   353 AA.
AC   O15975;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Guanine nucleotide-binding protein G(q), alpha subunit.
GN   SCGQA.
OS   Patinopecten yessoensis (Ezo giant scallop) (Yesso scallop).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Pectinoida;
OC   Pectinoidea; Pectinidae; Mizuhopecten.
OX   NCBI_TaxID=6573;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97435252; PubMed=9287291;
RA   Kojima D., Terakita A., Ishikawa T., Tsukahara Y., Maeda A.,
RA   Shichida Y.;
RT   "A novel Go-mediated phototransduction cascade in scallop visual
RT   cells.";
RL   J. Biol. Chem. 272:22979-22989(1997).
CC   -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC       INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC       SIGNALING SYSTEMS.
CC   -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC       THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC   -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC   -----------------------------------------------------------------------
```
```
CC   -----------------------------------------------------------------------
DR   EMBL; AB006456; BAA22219.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
KW   GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW   Palmitate; Lipoprotein.
FT   LIPID          3      3       PALMITATE (BY SIMILARITY).
FT   LIPID          4      4       PALMITATE (BY SIMILARITY).
FT   NP_BIND       40     47       GTP (BY SIMILARITY).
FT   NP_BIND      199    203       GTP (BY SIMILARITY).
FT   NP_BIND      268    271       GTP (BY SIMILARITY).
FT   MOD_RES      177    177       ADP-RIBOSYL(1) (BY ACTION OF CTX)
SQ   SEQUENCE   353 AA;  41237 MW;  5FDECC3C1C738B17 CRC64;

Query Match       82.3%;  Score 1530;  DB 1;  Length 353;
Best Local Similarity  82.4%;  Pred. No. 8.2e-108;
Matches 291;  Conservative 32;  Mismatches 30;  Indels 0;  Gaps 0;

Qy      7  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG  66
           |||||||||| :||| |||:| | :: ||||||||||||||||||||||||||||||| |
Db      1  MACCLSEEAKEQKRINCEIEKELRKAKRDARRELKLLLLGTGESGKSTFIKQMRIIHGTG  60

Qy     67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
           ||.|||||| |:|||||| || ::||||||:|| ::| .: | : : :||||:|:| :
Db     61  YSEEDKRGFIKIVYQNIFMAMHSMIRAMDTIKISFEVADNENAIMIRQVDYETVVTLDS  120

Qy    127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  186
           |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  YSEEDKRGFIKIVYQNIFMAM...
```

```
Db  121 QSVEAILSLWADAGIQECYDRREYQLTDSAKYYLDAVDRIAEPNYLPTLQDILRVRVPT 180

Qy  187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 246
            |:|||||||||| |:||||||||||||||||||||||||||||||||||||||||
Db  181 TGIIEYPFDLLDSIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN 240

Qy  247 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR 306
            |||||||||:|||||||||||||||||||||| :||||||| :::|||| ||
Db  241 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMHSHLVDYFPEFDGQKKDAQGAR 300

Qy  307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
            |||||:|||||||| ||||||||||||||||||||||||||||||||||||
Db  301 EFILRMFVDLNPDPDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353
```

Search completed: June 30, 2003, 16:16:29
Job time : 13 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:     June 30, 2003, 16:12:56 ; Search time 80 Seconds
                                      (without alignments)
                                      924.636 Million cell updates/sec

Title:         US-09-899-295-2
Perfect score: 1860
Sequence:    1 MTLESIMACCLSEEAKEARR.........VFAAVKDTILQLNLKEVNLV 359

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:   671580

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :    SPTREMBL_21:*
        1: sp_archea:*
        2: sp_bacteria:*
        3: sp_fungi1:*
        4: sp_human:*
        5: sp_invertebrate:*
        6: sp_mammal:*
        7: sp_mhc:*
        8: sp_organelle:*
        9: sp_phage:*
       10: sp_plant:*
       11: sp_rodent:*
       12: sp_virus:*
       13: sp_vertebrate:*
       14: sp_unclassified:*
       15: sp_rvirus:*
       16: sp_bacteriap:*
       17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1849 | 99.4 | 359 | 4 | Q9BZB9 | Q9bzb9 homo sapien |
| 2 | 1666 | 89.6 | 359 | 11 | Q91X95 | q91x95 mus musculu |
| 3 | 1544 | 83.0 | 353 | 5 | P91955 | p91955 limulus pol |
| 4 | 1532 | 82.4 | 353 | 5 | Q9U473 | q9u473 panulirus a |
| 5 | 1524 | 81.9 | 355 | 11 | Q8R2X9 | q8r2x9 mus musculu |
| 6 | 1513 | 81.3 | 355 | 5 | O17386 | o17386 caenorhabdi |
| 7 | 1504 | 80.9 | 353 | 5 | Q9NL92 | q9nl92 octopus vul |
| 8 | 1498 | 80.5 | 353 | 5 | Q8T6P8 | q8t6p8 mamestra br |
| 9 | 1439 | 77.4 | 360 | 5 | Q8WSU8 | q8wsu8 halocynthia |
| 10 | 1430 | 76.9 | 353 | 5 | Q9NFZ0 | q9nfz0 calliphora |
| 11 | 1291 | 69.4 | 303 | 5 | Q8T3G5 | q8t3g5 caenorhabdi |
| 12 | 1152.5 | 62.0 | 305 | 5 | Q9Y207 | q9y207 hydra magni |
| 13 | 1122.5 | 60.3 | 355 | 5 | Q9XZV4 | q9xzv4 geodia cydo |
| 14 | 1112 | 59.8 | 279 | 5 | Q9I7C8 | q9i7c8 drosophila |
| 15 | 1055 | 56.7 | 374 | 6 | Q9TU29 | q9tu29 oryctolagus |
| 16 | 902 | 48.5 | 353 | 3 | Q8TGA3 | q8tga3 leptosphaer |

ALIGNMENTS

| No. | Score | % | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 17 | 897.5 | 48.3 | 354 | 4 | Q9UGA4 | homo sapien |
| 18 | 897 | 48.2 | 353 | 3 | Q9C134 | tapesia yal |
| 19 | 896 | 48.2 | 353 | 3 | Q96VA7 | fusarium ox |
| 20 | 896 | 48.2 | 353 | 3 | Q96TI8 | gibberella |
| 21 | 895.5 | 48.1 | 354 | 13 | Q90847 | gallus gall |
| 22 | 893 | 48.0 | 353 | 3 | Q9NFA2 | roselinia |
| 23 | 892 | 48.0 | 353 | 3 | Q96VN7 | trichoderma |
| 24 | 887.5 | 47.7 | 354 | 13 | Q90846 | gallus gall |
| 25 | 885.5 | 47.6 | 354 | 4 | Q8TAN5 | homo sapien |
| 26 | 884.5 | 47.6 | 301 | 5 | Q9Y203 | ephydatia f |
| 27 | 883.5 | 47.5 | 354 | 4 | Q8TD72 | homo sapien |
| 28 | 883.5 | 47.5 | 355 | 11 | Q9DC51 | mus musculu |
| 29 | 882 | 47.4 | 305 | 5 | Q9y203 | ephydatia f |
| 30 | 881 | 47.4 | 469 | 5 | Q8WPA0 | halocynthia |
| 31 | 875 | 47.0 | 355 | 13 | Q9W6A4 | squalus aca |
| 32 | 874.5 | 47.0 | 354 | 3 | Q9DC51 | blumeria gr |
| 33 | 874 | 47.0 | 353 | 3 | Q9HFN1 | botrytis ci |
| 34 | 873 | 46.9 | 350 | 13 | Q9HEP9 | homo sapien |
| 35 | 870.5 | 46.8 | 354 | 5 | Q96C71 | octopus·vul |
| 36 | 868 | 46.7 | 355 | 4 | Q96C71 | homo sapien |
| 37 | 863.5 | 46.4 | 354 | 3 | Q9HF99 | schizophyll |
| 38 | 861.5 | 46.3 | 354 | 13 | O13011 | xenopus lae |
| 39 | 859 | 46.2 | 371 | 13 | Q8QGY8 | fugu rubrip |
| 40 | 858.5 | 46.2 | 350 | 13 | Q9I123 | ambystoma t |
| 41 | 858.5 | 46.2 | 354 | 13 | Q9C115 | pisolithus |
| 42 | 855.5 | 46.0 | 354 | 4 | Q8WP45 | halocynthia |
| 43 | 854.5 | 45.9 | 350 | 13 | Q90WX6 | brachydanio |
| 44 | 854.5 | 45.9 | 354 | 13 | Q9DG27 | gallus gall |
| 45 | 851.5 | 45.8 | 354 | 5 | Q8WSS1 | ciona intes |

RESULT 1
Q9BZB9

ID   Q9BZB9       PRELIMINARY;   PRT;   359 AA.
AC   Q9BZB9;
DT   01-JUN-2001 (TREMBLrel. 17, Created)
DT   01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE   G alpha q protein (Guanine nucleotide binding protein alpha q).
GN   GNAQ.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98366208; PubMed=9700850;
RA   Gabbeta J., Dhanasekaran N., Rao A.K.;
RT   "G alpha q cDNA sequence from human platelets.";
RL   Thromb. Res. 91:29-32(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gabbeta J., Dhanasekaran N., Rao A.K.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RT   "Homo sapiens guanine nucleotide binding protein (G protein) alpha q
RT   (GNAQ).";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF329284; AAG61117.1; -.
DR   EMBL; AF493896; AAM12610.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE   359 AA;  42142 MW;  6F69C4F617DFA7C7 CRC64;

Query Match     87.7%;     Score 1666;     DB 11;     Length 359;
Best Local Similarity     87.7%;     Pred. No. 5.7e-122;
Matches 315;     Conservative 29;     Mismatches 15;     Indels 0;     Gaps 0;

```
Qy    1  MTLESIMACCLSEAKEARRINDETERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR   60
         ||||| :|||||:|:||:|||| :|:|:|| ::||:|||||||||||||||||||||||
Db    1  MTLENLMACCLSDEVKESKRINAELEKQLRDKRDARRELKLLLLGTGESGKSTFIKQMR   60

Qy   61  IIHGAGYSEEDKRGFTKLVYQNIFTAMQAVRAMATLKILYKYEQNKANALLIREVDLEK  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  IIHGAGYSEEDKRGFTKLVYQNIFTAMQAVRAMATLKILYKYEQNKANALLIREVDLEK  120

Qy  121  VTTFEHQYVNAIKTLWSDPGVQECYDRRDPEFQLSDSAKYLTDVDRIATVGYLPTQQDVL  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  VTTFEHQYVNAIKTLWSDPGVQECYDRRDPEFQLSDSAKYLTDVDRIATVGYLPTQQDVL  180

Qy  181  RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
         ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db  181  RVRVPTTGIIEYPFDLENIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
```

RESULT 2
Q91X95
ID   Q91X95       PRELIMINARY;       PRT;       359 AA.
AC   Q91X95;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to guanine nucleotide binding protein, alpha 11.
GN   GNA11.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COLON;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011169; AAH11169.1; -.
DR   MGD; MGI:95766; Gna11.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   ProDom; PD000281; Gprotein_alpha; 1.
SQ   SEQUENCE   359 AA;   41930 MW;   7437912FF0BAC855 CRC64;

Query Match     99.4%;     Score 1849;     DB 4;     Length 359;
Best Local Similarity     99.4%;     Pred. No. 3e-136;
Matches 357;     Conservative 1;     Mismatches 1;     Indels 0;     Gaps 0;

```
Qy    1  MTLESIMACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMR   60
         |||||||||||||||||||||||||||| ::|||||||||||||||||||||||||||||
Db    1  MTLESIMACCLSEEAKEAARRINDEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMR   60

Qy   61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  IIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEK  120

Qy  121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  VSAFENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVL  180

Qy  181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  RVRVPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLV  240

Qy  241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  ESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQR  300

Qy  301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  DAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
```

RESULT 3
P91955
ID   P91955       PRELIMINARY;       PRT;       353 AA.
AC   P91955;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Gq protein alpha subunit.
OS   Limulus polyphemus (Atlantic horseshoe crab).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC   Limulidae; Limulus.
OX   NCBI_TaxID=6850;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENTRAL EYE;
RA   Munger S.D., Schremser-Berlin J.-L., Brink C.M., Battelle B.-A.;
RT   "Molecular and Immunological Characterization of a Gq Protein from
RT   Ventral and Lateral Eye of the horseshoe Crab Limulus polyphemus.";
RL   Invert. Neurosci. 0:0-0(1997).
DR   EMBL; U88586; AAB48510.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE   353 AA;   41497 MW;   C1C57783B3D2D516 CRC64;

Query Match     83.0%;     Score 1544;     DB 5;     Length 353;
Best Local Similarity     83.0%;     Pred. No. 1.8e-112;
Matches 293;     Conservative 28;     Mismatches 32;     Indels 0;     Gaps 0;

```
Qy    7  MACCLSEEAKEAARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG   66
         ||||||| :||: :||:||||| :.|||||||||||||||||||||||||||||||:||:|
Db    1  MACCLSEEGKEQKRINQEIERQLRDKRDARRELKLLLLGTGESGKSTFIKQMRIHGQG    60

Qy   67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
         |||:||| | ||||||| ||||:|.:||:|:||||||| :: |  :.||:.||  ::
Db   61  YSDDDKSYIKLVYQNIIMAMQSMNKAMEMLKISYKDRNNIENAELVLSVDYETVTFDS   120

Qy  127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  186
         |||:|||||| ||||||||||||||||.||:||||||:||:|.|  ||||||:|||||||
Db  121  PYVEAIKSLWVDPGIQECYDRRREYQLTDSAKYYLNDIDRIAVPNYLPTQQDILRVRVPT  180

Qy  187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
         |||||||| | :|||||||||||||||||||||||||||||||||||||||||||||||
Db  181  TGIIEYPFILDSIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy  247  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPKKDAVQGR  306
         |||||||| :|||||||| |||||||||||||||||| ||||||||||||||||||||||
Db  241  RMEESKALFKTIITYPWFLNSSVILFLNKKDLLEEKIMFSHLVDYFPEYDGPKKDAVQGR  300

Qy  307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV  359
         ||||||||||||||:||||||||||||||||||||||||||| |||||||||
Db  301  EFILKMFVDLNPDSEKIIYSHFTCATDTENIRFVFAAVKDTILHSHLVDYFPEFDGPQR  353
```

RESULT 4
Q9U473
ID   Q9U473       PRELIMINARY;       PRT;       353 AA.
AC   Q9U473;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```
RESULT 5
ID   Q8R2X9
AC   Q8R2X9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Guanine nucleotide binding protein, alpha 14.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC027015; AAH27015.1; -.
SQ   SEQUENCE 355 AA; 41527 MW; D34B39ACD179AE82 CRC64;
```

```
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Gq/11 protein alpha subunit.
OS   Panulirus argus (Spiny lobster)
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;
OC   Palinuroidea; Palinuridae; Panulirus.
OX   NCBI_TaxID=6737;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OLFACTORY ORGAN;
RX   MEDLINE=20347127; PubMed=10781594;
RA   Munger S.D., Gleeson R.A., Aldrich H.C., Rust N.C., Ache B.W.,
RA   Greenberg R.M.;
RT   "Characterization of a phosphoinositide-mediated Odor transduction
RT   pathway reveals plasma membrane localization of an inositol 1,4,5-
RT   trisphosphate receptor in lobster olfactory receptor neurons.";
RL   J. Biol. Chem. 275:20450-20457(2000).
DR   EMBL; AF201328; AAF19378.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE 353 AA; 41478 MW; FD9B551F66327BAA CRC64;

PRELIMINARY;        PRT;        355 AA.

Query Match    82.4%;  Score 1532;  DB 5;  Length 353;
Best Local Similarity  83.0%;  Pred. No. 1.6e-111;
Matches 293;  Conservative 24;  Mismatches 36;  Indels 0;  Gaps 0;

Qy    7 MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLLGTGESGKSTFIKQMRIIHGSG 66
         |||||||||||| |||| |||| :||||||||||||||||||||||||||||||||||:|
Db    1 MACCLSEEAKEQKRINQEIEROLRKDKRDARRELKLILLLGTGESGKSTFIKQMRIIHGAG 60

Qy   67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENY 126
         |||||||||:||||:||||:|||:||||| |||::|: | : ||||| ||::||:|:|::|
Db   61 YSDEDKRGFIKLVFQNIFMAMQSMIRAMDLLQISYGDSANIEHADLVRGVDVESVTFEE   120

Qy  127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQODVLRVRVPT  186
         |:| |:|||| ||||: |||||||||| ||:|||| ||| | || | ||:|:||||:|:|
Db  121 PVVTAMKSLWQDPGIQHCYDRRREYQLTDSAKYYLTLDRIAATDYSTILQDILRVRAPT   180

Qy  187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
         |||||||||:  :|||||||||||||||||||||||||||||||||||||||||:||||
Db  181 TGIIEYPFDLEIIRFRMVDVGGQRSERRKWIHCFENVTSIMFIVALSEYDQVLVEDNENRM 246

Qy  247 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
         ||||||||:|||||||||:|||||||||||||||||||||||||||||||||:: ||:||
Db  241 RMEESKALFKTIITYPWFQHSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPRKDAIAAR  300

Qy  307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         ||||:|||:||:| :||||||||||||||||||||||||||||||||||||||
Db  301 EFILRMFVELNDPPEKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 353
```

```
RESULT 6
ID   Q17386
AC   Q17386; O02546;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   EGL-30 (HETEROTRIMERIC G protein alpha subunit).
GN   EGL-30 OR M01D7.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   STRAIN=N2;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RL   Neuron 16:999-1009(1996).
RT   "Mutations in a C. elegans Ggalpha gene disrupt movement, egg laying,
RT   and viability.";
RA   Brundage L., Avery L., Katz A., Kim U.J., Mendel J.E., Sternberg P.W.,
RA   Simon M.I.;
RX   MEDLINE=96221161; PubMed=8630258;
RP   SEQUENCE FROM N.A.
RN   [1]
OX   NCBI_TaxID=6239;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OS   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
GN   Caenorhabditis elegans.
DE   EGL-30 (HETEROTRIMERIC G protein alpha subunit).
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
AC   Q17386; O02546;
ID   Q17386
```

```
PRELIMINARY;        PRT;        355 AA.

Query Match    81.9%;  Score 1524;  DB 11;  Length 355;
Best Local Similarity  81.58%;  Pred. No. 6.6e-111;
Matches 286;  Conservative 36;  Mismatches 29;  Indels 0;  Gaps 0;

Qy    9 CCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLLGTGESGKSTFIKQMRIIHGSGYS 68
         ||||  ||| :||| :|||| :|||||:||||||||||||||||||||||||||||||| ||
Db    5 CCLSAEEKESQRISAEIEROLRRDKRDDARRELKLILLLGTGESGKSTFIKQMRIIHGSYS 64

Qy   69 DEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEVDKVTALSRDQ 128
         |||:||||||||:||||:|||||||:||||:|:| ||:|   |:||  |||:|:|:: : |
Db   65 DEDRGFTKLVYQNIFTAMQAMIRAMDTLRIQYMCEQNKENAQIIREVEVDKVTALSRDQ  124

Qy  129 VDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQODVLRVRVPTTG  188
         |:| ||||||||:||||||||||||:|:||||| |||:||:|:||||:|:||||||||||
Db  125 VAAIKQLWLDPGIQHCYDRRREYQLTDSAKYYLTIDIERIAMPSFVPTQODVLRVRVPTTG 184

Qy  189 IIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRM 248
         |||||||| ::|||||||||||||||||||||||||||| |||||||||||||| ||||:
Db  185 IIEYPFDLENIIRFRMVDVGGQRSERRKWIHCFESVTSIIFLVALSEYDQVLAECDNENRM 244

Qy  249 EESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREF 308
         |||||||:||||||||:||||||||||||||||||||||||||||||||||: ||:|||:|
Db  245 EESKALFRTIITYPWFLNSSVILFLNKKDLLEEKIMYSHLSIYFPEYTGPKQDVKAARDF 304

Qy  309 ILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
         || ::||||::| :||||||||||||||||||||||||||||||:|:||||
Db  305 ILKIYQDQNPDKEKVIYSHFTCATDTENIRFVFAAVKDTILQLNLREFNLV 355
```

```
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN-BRISTOL N2;
RA   Waterston R.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Cuppen E., Jansen G., Plasterk R.H.A.;
RT   "Interaction analysis of the complete G-alpha subfamily of
RT   heterotrimeric G proteins from Caenorhabditis elegans.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U56864; AAB04059.1; -.
DR   EMBL; AF003739; AAB58071.1; -.
DR   EMBL; AY008139; AAG32092.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE  355 AA;  41865 MW;  FEA38B01C2E1355C CRC64;

Query Match           81.3%;
Best Local Similarity  82.0%;   Score 1513;  DB 5;  Length 355;
Matches  291;  Conservative  26;  Mismatches  36;  Indels  2;  Gaps  1;

Qy    7  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMRIIHGSG   66
            |||||||| ||||| ||| ||||||||||||||||||:|||||||||||||||||||||
Db    1  MACCLSEEAREQKRINQEIEKQLQRDKRNARRELKLILLGTGESGKSTFIKQMRIIHGQG   60

Qy   67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPY--KYEHNKAHAQLVREVDVEKVSAF  124
            |:||||||| |||||||| |:|||||||||||||    ||| || | ||||||||||||
Db   61  YSEEDKRAHIRLVYQNVFMAIQSMIRAMDTLDIKFGNESEELQEKAAVVREVDFESVTSF  120

Qy  125  ENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRV  184
            | |||| | |||||:||||||:|||||||||:||| ||||:||||||||||||||||||
Db  121  EEPYSYIKELWEDSGIQECYDRRREYQLTDSAKYYLSDLRRLAVPDYLPTEQDILRVRV  180

Qy  125  ENPYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRV  184
Db  121  EEPYSYIKELWEDSGIQECYDRRREYQLTDSAKYYLSDLRRLAVPDYLPTEQDILRVRV  180

Qy  185  PTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDN  244
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |||
Db  181  PTTGIIEYPFDLEQIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVECDN  240

Qy  245  ENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQA  304
            ||||||||||||||||||||| ||||||||||||||||| |||| ||||||||| ||
Db  241  ENRMEESKALFRTIITYPWFTNSSVILFLNKKDLLEEKILYSHLADYFPEYDGPPRDPIA  300

Qy  305  AREFILKMFVDLNPDSDKIIYSHFTCATDTENIREVFAAVKDTILQLNLKEYNLV  359
            ||||||||||||||| | |||||||||||||||| ||||||||||| ||||||||
Db  301  AREFILKMFVDLNPDADKIIYSHFTCATDTENIRFVFAAVKDTILQHLNLKEYNLV  355


RESULT 7
ID   Q9NL92           PRELIMINARY;     PRT;   353 AA.
AC   Q9NL92;
     Q9NL92
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   G protein a subunit q class.
OVGAQ.
OS   Octopus vulgaris (Octopus).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC   Incirrata; Octopodidae; Octopus.
OX   NCBI_TaxID=6645;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EYE;
RA   Iwasa T., Yanai T., Nakagawa M., Kikkawa S., Obata S., Usukura J.,
RA   Tsuda M.;
RT   "G protein alpha subunit genes in octopus photoreceptor cells.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AB025782; BAA93638.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE  353 AA;  40996 MW;  1AE03CB8CE28ED4E CRC64;

Query Match           80.9%;
Best Local Similarity  80.2%;   Score 1504;  DB 5;  Length 353;
Matches  283;  Conservative  31;  Mismatches  39;  Indels  0;  Gaps  0;

Qy    7  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMRIIHGSG   66
            ||||||||| || ||| |||| |||||||||||||||||||||||||||||||||||:|
Db    1  MACCLSEEAKEQKRINQEIERQLRRDKRDARRELKLILLGTGESGKSTFIKQMRIIHGAG   60

Qy   67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
            |:|||||| |:||||||| |:| : | |:||| :| |          |:: :    | :
Db   61  YSEEDKRGFKIVYQNIFSAIQTLIAAMETLSLEYKITSGNNENAEFIDSIDADSADTFDQ  120

Qy  127  PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPT  186
            :|||||||||.| | ||||||||||||:||:|||| | :||  |   |||||:|||||||
Db  121  SHVDAIKSLWTDEGMQECYDRRREYQLTDSAKYYLDDTERLHEPGHSPTLQDILRVRVPT  180

Qy  187  TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  246
            ||||||||||| | |||.|||||||||||||||||||||||||||||||||||||||||
Db  181  TGIIEYPFDLDSTILRMADVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNEN  240

Qy  247  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAR  306
            ||||||||||||||||||||||||||||||||||||||| ||| ||||||| |.|||||
Db  241  RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMTSHLADYFPEYDYDGPKKDAQAAR  300

Qy  307  EFILKMFVDLNPDSDKIIYSHFTCATDTENIREVFAAVKDTILQLNLKEYNLV  359
            ||||||||||||| |||||||||||||||||| |||||||||||| ||||||
Db  301  EFILKMFVDLNPDPDKIIYSHFTCATDTENIRFVFAAVKDTILQHLNLKEYNLV  353


RESULT 8
ID   Q8T6P8           PRELIMINARY;     PRT;   353 AA.
AC   Q8T6P8;
     Q8T6P8
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative Gq protein alpha subunit.
OS   Mamestra brassicae.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Hadeninae; Mamestra.
OX   NCBI_TaxID=55057;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Jacquin-Joly E., Francois M.-C., Burnet M., Lucas P., Bourrat F.,
RA   Maida R.;
RT   "Expression pattern in the antennae of a newly isolated lepidopteran
RT   Gq protein.";
RL   Eur. J. Biochem. 269:0-0(2002).
DR   EMBL; AF448447; AAL93221.1; -.
SQ   SEQUENCE  353 AA;  41390 MW;  8137E9A1EB7E87AB CRC64;

Query Match           80.5%;
Best Local Similarity  80.8%;   Score 1498;  DB 5;  Length 353;
Matches  282;  Conservative  32;  Mismatches  38;  Indels  0;  Gaps  0;

Qy    7  MACCLSEEAKEARRINDEIERHVRRDKRDARRELKLILLGTGESGKSTFIKQMRIIHGSG   66
            |||||||||||.| ||| |||| |:|:||||||||||||||||||||||||||||||||
Db    1  MACCLSEEAKEQKRINQEIERQLRKDRDARRELKLILLGTGESGKSTFIKQMRIIHGSG   60

Qy   67  YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN  126
            |:|:|||||.|||||||| |||:||||||.|| :|          | | ::  :|  : :
Db   61  YSDDDKRGFIKLVYQNIFMAMQSMIRAMELLTIQYGNPSNSEKAELISSIDFESVTFES  120
```

```
Qy  127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVPT 186
        ||||:|||:||||| |||||||| ||||: |||| :||||| :|||| ||:|||||||| |
Db  121 PYVEAIKALWADAGIQECYDRREYQLTDSAKYLQEIDRVAAPNYLPTEQDILRVPLT 180

Qy  187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDVLVESDNEN 246
        |||||||||:: ||||||||||||||||||||||||||||:||||||||: ||:| :||
Db  181 TGIIEYPFDLEEIRFRMVDVGGQRSERRKWIHCFENVTSIIFLVALSEYDQILFESENEN 240

Qy  247 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQPRDAQAAR 306
        ||||||||| :||||||||:||||||||||||||||||||||||||:|||||| ||||:||
Db  241 RMEESKALFKIIITYPWFQHSSVILFLNKKDLLEEKIMYSHLVDFPPEYDGSPQRDANTAR 300

Qy  307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
        ||||: |||||||::|||||||||||||||||:|| :||||||||:||:|::|:|| ||
Db  301 EFILRTFVDLNPDAEKIIYSHFTCATDTENIKLVFCAVKDTIMQSALKEFNL 352


RESULT 9

ID   Q8WSU8        PRELIMINARY;      PRT;    360 AA.
AC   Q8WSU8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein alpha subunit q class.
GN   HRGQ.
OS   Halocynthia roretzi (Sea squirt).
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC   Stolidobranchia; Pyuridae; Halocynthia.
OX   NCBI_TaxID=7729;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Iwasa T., Kanehara K., Watari A., Ohkuma M., Tsuda M.;
RT   "Cloning and expression patterns of G proteins in the larvae of
RT   Halocynthia roretzi.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB047084; BAB79199.1; -.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE  360 AA;  42571 MW;  A4ECFB6AE3C56DC9 CRC64;

Query Match      77.4%;   Score 1439;   DB 5;   Length 360;
Best Local Similarity   78.3%;   Pred. No. 2.9e-104;
Matches 282;   Conservative 27;   Mismatches 47;   Indels 4;   Gaps 2;

Qy    3 LESIMACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRII 62
        ||:||  ||:||| : ||||| |||||  ||||||||||||:|||||||||||||||||
Db    2 LEIMEFCCKSEEQRERAKRINAEIERELRHKRDARRELKLLLLGTGESGKSTFIKQMRII 61

Qy   63 HGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVS 122
        ||.|||| ||| :||| :||| ||| ||||||:||:| ||::: :  : :|| ||| |:
Db   62 HGKGYSDNDRREFCKLVHQNILTSMQNMIRAMETLHIQYSDPQNEQNSYDFLELDVERV- 120

Qy  123 AFEN---PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDV 179
        | :|   || ||:|: || : | |||||||||||||| || ||:||: | ||||||:||
Db  121 AEQNGPEYCTAIEALWTDKGTHDCYDRREYQLSDSTSYLNSLDRIRMPDYLPTEQDV 180

Qy  180 LRVRVPPTTGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVL 239
        ||||||| |||::||||||:: :||||||||||||||||||||||||||||||||||||||
Db  181 LRVRVPPTGISEYPFDLDSIIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVL 240

Qy  240 VESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQ 299
        |||:||||||||||||||| ||||||||||||||||||||||||||||||| |||:|||
Db  241 VESNNENRMEESKALFRTIIMYPWFQNSSVILFLNKKDLLEEKIMYSHLADYFSDVDGPT 300

Qy  300 RDAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV 359
        :| :|||||| : ||:|||||::||:|||||||||||||||||||||||||:|| ||||||
Db  301 QDANAAREFILSRFVELNPDADKIVYSHFTCATDTENIRFVFAAVKDTIMQCNLREYNLV 360
```

```
RESULT 10

ID   Q9NFZ0        PRELIMINARY;      PRT;    353 AA.
AC   Q9NFZ0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Guanine nucleotide-binding protein alpha subunit.
OS   Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=7373;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulz S., Huber A., Schwab K., Paulsen R.;
RT   "A novel Ggamma subunit isolated from Drosophila constitutes a visual G
RT   protein gamma subunit of the fly compound eye.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ250443; CAB76453.1; -.
DR   HSSP; P10824; 1BOF.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE  353 AA;  41295 MW;  FFF799774C02072E CRC64;

Query Match      76.9%;   Score 1430;   DB 5;   Length 353;
Best Local Similarity   76.1%;   Pred. No. 1.4e-103;
Matches 268;   Conservative 39;   Mismatches 45;   Indels 0;   Gaps 0;

Qy    7 MACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
        | ||||| |:::|||::|||: |||||||||||||||||||||||||||||||||||||
Db    1 MECCLSEAKEQKRINQEIERQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 60

Qy   67 YSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
        ||: |||| ||| :||| ||||:||:|||:|:   |   :  : :| :::||: |:| :
Db   61 YSDDDKRGYIKLVFQNINFAMQSMIKAMDMLRISYGARDHAEBLAELVMSIDYETVTSFED 120

Qy  127 PYVDAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVPT 186
        || :|||:|| : ||||||||||||||:||||||||:|:|:  | |||||:|:||||||
Db  121 PYLSAIKILWADTGIQECYDRREYQLTDSAKYYLSDLERIEQADYLPSEQDILRARVPT 180

Qy  187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDVLVESDNEN 246
        |||||||||| : ||||||||||||||||||||||||||||:||||||||: |:||||||
Db  181 TGIIEYPFDLGIVFRMVDVGGQRSERRKWIHCFENVTSIIFLVALSEYDQILFESDNEN 240

Qy  247 RMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQPRDAQAAR 306
        |||||||||||||:||||||||||||||||||||||||||||||||||||||| | ||:|:
Db  241 RMEESKALFRTIIYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPKCDHAAAK 300

Qy  307 EFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNL 358
        : |:||: :  ||||:  | |||||:|||||||:||| ||||||:|| |:||| |
Db  301 QFVLKKYLAAANPDPERQCTSHFTTATDTENIKLVFCAVKDTIMQNALKRFNL 352


RESULT 11

ID   Q8T3G5        PRELIMINARY;      PRT;    303 AA.
AC   Q8T3G5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   C. elegans EGL-30 protein (corresponding sequence M01D7.7b).
GN   EGL-30.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditina;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
```

RESULT 12

```
ID   Q9Y207
AC   Q9Y207;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   G protein a subunit 3 (Fragment).
OS   Hydra magnipapillata (Hydra).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida;
OC   Hydridae; Hydra.
OX   NCBI_TaxID=6085;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99246375; PubMed=10229568;
RA   Suga H., Koyanagi M., Hoshiyama D., Ono K., Iwabe N., Kuma K.,
RA   Miyata T.;
RT   "Extensive gene duplication in the early evolution of animals before
RT   the parazoan-eumetazoan split demonstrated by G proteins and protein
RT   tyrosine kinases from sponge and hydra.";
RL   J. Mol. Evol. 48:646-653(1999).
DR   EMBL; AB006541; BAA81695.1; -.
DR   HSSP; P10824; 1B0F.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   305 AA;  35491 MW;  2639884D4385869 CRC64;
```

```
Query Match      62.0%;  Score 1152.5;  DB 5;  Length 305;
Best Local Similarity  72.5%;  Pred. No. 5.1e-82;
Matches  221;  Conservative  31;  Mismatches  48;  Indels  5;  Gaps  1;
```

```
Qy   60 RIIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHHKAHAQLVREVDVE 119
         |||||||||||||||| |||:||||:|||||||||||||||||||:|||:||||||||||:
Db    1 RIIHGQGYTEDDRRGYTNLVFLNIYQAMQALTRAMRNLKISYSNPANEENARLILDVDLS 60

Qy  120 K----VSAFENPYVDAIKSLWNDPGIQOECYDRRREYQLSDSTKYYLNDLDRVADPAYLP 174
         |    ||||||||||||||||||||||||||||||||| |||:||||||||||:|::|||
Db   61 EDKSRTITLSPOYASAIESLWKDSGIQEEVDRREYQLSDSAKYYLSDLRRICANFVP 120

Qy  175 TQODVLRVRVPPTGIIEYPFDLQSVIFRMVDVGGQQRSERRKWIHCFENVTSIMFLVALSE 234
         || |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db  121 TMQDVLRARAPTTGIIEYPFDLDTIIFRMVDVGGQQRSERRKWIHCFENVTSIMFLVALSE 180

Qy  235 YDQVLVESDNENRMEESKALFRTIITYPWFQNSSVILFLFNKKDLLEEKIMYSHLVDYFPPE 294
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 YDQVLFESQSENRMDESKALFRTIITYPWFLQSSIILFLNKTDLLEEKIQKSDLQTFPE 240

Qy  295 YDGPQRDAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLK 354
         |||||||||||||||||||||||||:|||||||||||||||||||||||:||||||||||
Db  241 YDGAKGDAKAAKEFILKMFVDLNPDTDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLK 300

Qy  355 EYNLV 359
         |||||
Db  301 EYNLV 305
```

RESULT 13

```
ID   Q9XZV4
AC   Q9XZV4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Gq protein, alpha subunit.
OS   Geodia cydonium (Sponge).
OC   Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC   Astrophorida; Geodiidae; Geodia.
OX   NCBI_TaxID=6047;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98119577; PubMed=9459489;
RA   Seack J., Kruse M., Mueller W.E.G.;
RT   "Evolutionary analysis of G-protein alpha and beta-subunits from the
RT   sponge Geodia cydonium: cloning of alpha- and beta-subunits in early
RT   metazoans.";
RL   Biochim. Biophys. Acta 1401:93-103(1998).
DR   EMBL; Y14248; CAB43527.1; -.
DR   HSSP; P04896; 1AZT.
DR   InterPro; IPR001019; Gprotein_alpha.
DR   Pfam; PF00503; G-alpha; 1.
DR   PRINTS; PR00318; GPROTEINA.
DR   ProDom; PD000281; Gprotein_alpha; 1.
DR   SMART; SM00275; G-alpha; 1.
SQ   SEQUENCE   355 AA;  41363 MW;  6828C29643F2CE91 CRC64;
```

```
Query Match      60.3%;  Score 1122.5;  DB 5;  Length 355;
Best Local Similarity  61.4%;  Pred. No. 1.4e-79;
Matches  215;  Conservative  51;  Mismatches  83;  Indels  1;  Gaps  1;
```

```
Qy    7 MACCLSEEAKEARRINDEIEHVRDKRDARRELKLLLGTGESGKSTFIKQMRIIHGSG 66
         :|:|||| ||::||||||||:|||||:||||:|||::|||||||||||||||||||||
Db    1 MSCLLSEEERLQKRINTRINRELQRDHKDAKKEIKLLLLGTGESGKSTFIKQMRIIHGKG 60
```

Also appearing in the earlier column:

```
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99096613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Gattung S., Goela D.;
RT   "The sequence of C. elegans cosmid M01D7.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF003739; AAM15593.1; -.
SQ   SEQUENCE   303 AA;  35807 MW;  12B5A76956116D28 CRC64;
```

```
Query Match      69.4%;  Score 1291;  DB 5;  Length 303;
Best Local Similarity  81.8%;  Pred. No. 8e-93;
Matches  248;  Conservative  20;  Mismatches  33;  Indels  2;  Gaps  1;
```

```
Qy   59 MRIHGSGYSDEDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPY--KYEHHKAHAQLVREV 116
         ||||||||||||:||||:|||||||:|:|||||||||||||||    ::|||||||:|||
Db    1 MRIIHGQGYSEEDKRAHIRLVYQNVFMAIQSMIRAMDTLDIKFGNESEELQEKAAVREV 60

Qy  117 DVEKVSAFENPYVDAIKSLWNDPGIQOECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQ 176
         | |::| ||||||||||||||||||||| ||||||||| |:|:|||||||||||| :||:
Db   61 DFESVTSFEEPYVSVIKELWEDSGIQEECVDRRREYQLTDSAKYYLSDLRRLAVPDYLPTE 120

Qy  177 QDVLRVRVPPTGIIEYPFDLQSVIFRMVDVGGQQRSERRKWIHCFENVTSIMFLVALSEYD 236
         ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||:|
Db  121 QDILRVRVPPTGIIEYPFDLEQIFRMVDVGGQQRSERRKWIHCFENVTSIMFLVALSEYD 180

Qy  237 QVLVESDNENRMEESKALFRTIITYPWFQNSSVILFLFNKKDLLEEKIMYSHLVDYFPEYD 296
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 QVLVECDNENRMEESKALFRTIITYPWFTNSSVILFLNKKDLLEEKILYSHLADYFPEYD 240

Qy  297 GPQRDAQAAREFILKMFVDLNPDSDKIIYSHFTCATDTENIRFVFAAVKDTIILQLNLKEY 356
         ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db  241 GPPRDPIAAREFILKMFVDLNPDADKIIYSHFTCATDTENIRFVFAAVKDIILQHNLKEY 300

Qy  357 NLV 359
         |||
Db  301 NLV 303
```

```
Qy    67 YSDEDKRGFTKLYQNIFFAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFEN 126
         ||:|:|||  ||||||||||||||||| ||||| || |||| ||||||| | |||||||
Db    61 YSKQDCLEYKNLVFRNILMSMHSMLQATAELKIAYIDPPDAQRHVQLLMALRPETAQSLGG 120

Qy   127 PYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLLDRVADPAYLFPTQQDVLRVRPT 186
         :|||||||||||  ||||||||||||||||||:|||:|    || ||||||:|||||||||
Db   121 ETCEAIRKLWQDAGVOECYQRRNEYQLSDSTKYYLDDLPRISSNDYVPTTQDVLRVRPT 180

Qy   187 TGIIEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNE- 245
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGINEYPFTINKIFKMVDVGGQRSERRKWIHCFDHVTSVMFLVAISEYDQILVEADSRV 240

Qy   246 NRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQOA 305
         |:| |||||| |||:||||||||||||||||||||||||| ||:|  |  || |  | :|
Db   241 NRMVESLHLFNTIISYPWFNKSSIILFLNKKDLLEEKVMHSHLIDYFEEYDGPKCDHVSA 300

Qy   306 REFILKMFVDLNPDSDKIIYSHFTCATDTENTRRVFAAVKDTIIQLNLKE 355
         |||||||||||||| ||:|| ||  ||||||||||| || | ::|  | |
Db   301 RESIAKMFISINDMRSADIYPHFTCATDTENIKPVFDVVKNHILQQHITE 350
```

RESULT 14

```
Q9I7C8
ID   Q9I7C8     PRELIMINARY;    PRT;   279 AA.
AC   Q9I7C8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein alpha 49B.
GN   G-ALPHA-49B OR CG17759.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
EMBL; AE003821; AAG22276.1; -.
HSSP; P10824; 1BOF.
FlyBase; FBgn0004435; G-alpha-49B.
InterPro; IPR001019; Gprotein_alpha.
Pfam; PF00503; G-alpha; 1.
PRINTS; PR00318; GPROTEINA.
ProDom; PD000281; Gprotein_alpha; 1.
SMART; SM00275; G-alpha; 1.
SEQUENCE 279 AA;  32844 MW;  A9609C515FF24796 CRC64;
```

```
Query Match             59.8%;  Score 1112;  DB 5;  Length 279;
Best Local Similarity   76.3%;          Pred. No. 6.5e-79;
Matches  213; Conservative  27; Mismatches  19; Indels  20; Gaps  2;
```

```
Qy     7 MACCLSEEAKEARRINDEIERHVRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 66
         |||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||
Db     1 MECCLSEEAKEQKRINQEIEKQLRDDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSG 60

Qy    67 YSDEDKRGFTKLYQNIFFAMQAMIRAMDTLKIPY-KYEHNKAHAQLVREVDVEKVSAFE 125
         ||||||||| || |||||||||| ||:|| ||   : ||
Db    61 YSDEDKRGYIKLVFRNILMSMHSMLQATAELKIAYIDPPDAQRHVQLLMALRPE------ 101
```

RESULT 15

```
Q9TU29
ID   Q9TU29     PRELIMINARY;    PRT;   374 AA.
AC   Q9TU29;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
DE   Galpha16.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20035837; PubMed=10571060;
RA   Feild J.A., Foley J.J., Testa T.T., Nuthulaganti P., Ellis C.,
RA   Sarau H.M., Ames R.S.;
RT   "Cloning and characterization of a rabbit ortholog of human Galpha16
RT   and mouse G(alpha)15.";
RL   FEBS Lett. 460:53-56(1999).
EMBL; AF169627; AAF06740.1; -.
HSSP; P10824; 1AS3.
InterPro; IPR001019; Gprotein_alpha.
Pfam; PF00503; G-alpha; 1.
PRINTS; PR00318; GPROTEINA.
ProDom; PD000281; Gprotein_alpha; 1.
SMART; SM00275; G-alpha; 1.
SEQUENCE 374 AA;  42933 MW;  630B2920A2080CA1E CRC64;
```

```
Query Match             56.7%;  Score 1055;  DB 6;  Length 374;
Best Local Similarity   55.8%;          Pred. No. 2.7e-74;
Matches  202; Conservative  62; Mismatches  86; Indels  12; Gaps  2;
```

```
Qy   126 NPYVDAIKSLWNDPGIQECYDRREYQLSDSTKYYLNDLLDRVADPAYLFPTQQDVLRVRVP 185
         ||||||||||||||||||||||||||||||| ||:||:|||||||||||||||:|||| ||
Db    61 YSDEDKRGYIKLVFRNILMSMHSMLQATAELKIAYIDPPDAQRHVQLLMALRPE------- 101

Qy   186 TTGIIEYPFDLEEIRFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQILFESDNE 245
         ||||||||||:|:|||||||||||||||||||||:|||||||||:||||||||| ||||:
Db   162 TTGIIEYPFTINKIFKMVDVGGQRSERRKWIHCFDHVTSVMFLVAISEYDQILVEADSRV 221

Qy   246 NRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIM 284
         |:|:|:|||||||| |||||:|:  :| :|| ||:||||
Db   222 NRLEESKALFHTIITFEWFKNASIILFLNKMDVLEEKIM 260
```

```
Qy    10 CLSEEAKEARRINDEIERHVRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSD 69
```

```
Db    13 CLSEDEKAAAARVDQEITRLLLEHRRQVRGELKLILLGTGESGKSTFIKQMRIIHGAYSE  72
         ||||  |  |    |:  ||  |     |:    |::||::  ::|:::   : |:  :||:
Qy    70 EDKRGFTKLVYQNIFTAMQAMIRAMDTLKIPYKYEHNKAHAQLVREVDVEKVSAFENPYV 129
         |::||| |  |  |::| |    |   | | ||     |  || |  |::|   |  |  |
Db    73 EDRKGFPLVFQNIFLSVQAIIEAMDRLQIPYSRPESKLHASLVMSQDPYKVNTFETRYA 132

Qy   130 DAIKSLWNDPGIQECYDRRREYQLSDSTKYYLNDLDRVADPAYLPTQQDVLRVRVPTTGI 189
         |::||| |  | ||  | |  |:||||  || |||::| ||  | |||  |:| | ||||
Db   133 LAVQSLWRDAGVRACYERRREFHLLDSAVYLSHLERIAEEGYVPTAQDVLRSRMPTTGI 192

Qy   190 IEYPFDLQSVIFRMVDVGGQRSERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRME 249
         ||  | |  |:   ||:|||:|||||||||||||||| : |  |||||::||:|||:|:
Db   193 NEYCFSVQKTNLRIVDVGGQKSERRKWIHCFEDVTALIFLASLSEYDQCLEENGQENRMQ 252

Qy   250 ESKALFRTIITYPWFQNSSVILFLNKKDLLEEKIMYSHLVDYFPEYDGPQRDAQAAREFI 309
         ||| || :|  ||| | :| | :||:||  | :|  || |||||||||||:|   ||:||
Db   253 ESLALFGTVLALPWFRATSVILFLNKTDILEDKVRTSHLATYFPGFRGPPQDPEAAKRFI 312

Qy   310 LKMFVDL-------NPDS-------DKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
         |  |     ||  : :|:||||||:|||||:|||| || | | |  :| |
Db   313 LELYTRVYAGAAAGPDGASKGPRSRRLFSHYTCATDTQNIRKVFKDVRDSVLARYLDEIN 372

Qy   358 LV 359
         |:
Db   373 LL 374
```